

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:14 ; Search time 80.7044 Seconds  
(without alignments)  
2741.199 Million cell updates/sec

Title: US-10-017-479A-2

Perfect score: 2994

Sequence: 1 MEIEIGBQPPVKCNFFA.....NSFPWAQIYAAALGNKTH 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	2994	100.0	572	4	ABB60315	Abb60315 Drosophil	
2	2994	100.0	572	4	ABB66903	Abb66903 Drosophil	
3	2994	100.0	572	5	ABB79611	Abb79611 Drosophil	
4	2994	100.0	572	8	ADP64789	ADP64789 Drosophil	
5	2986	99.7	572	6	ABR40099	ABr40099 dIndyaa s	
6	1327.5	44.3	908	4	ABB61068	ABB61068 Drosophil	
7	1042.5	34.8	587	7	AAE38766	AAe38766 Rat 69624	
8	1039	34.7	581	8	ADP64799	ADP64799 Zebrafish	
9	1026	34.3	592	6	ABB82950	ABB82950 Human SLC	
10	1026	34.3	592	6	ABB82951	ABB82951 Human SLC	
11	1026	34.3	592	6	ABR40100	ABr40100 Human ren	
12	1026	34.3	592	6	ABR57025	ABr57025 Human NaD	
13	1026	34.3	592	7	ADP63755	ADe63755 Human Pro	
14	1010.5	33.8	619	8	ADG16995	Adg16995 African c	
15	1005	33.6	576	8	ADH42443	ADh42443 Novel hum	
16	1004	33.5	568	6	ABR57023	ABr57023 Human TCH	
17	1004	33.5	568	7	AAE38764	AAe38764 Human 696	
18	1004	33.5	568	7	ADK51052	ADk51052 Human NOV	
19	1004	33.5	568	8	ADH42441	ADh42441 Novel hum	
20	1004	33.5	568	8	ADP64793	ADP64793 Human Na+	
21	1002	33.5	568	5	ABU65062	ABu65062 Human NOV	
22	1002	33.5	568	6	ABR40097	ABr40097 Human sod	
23	1002	33.5	568	7	ADK51050	ADk51050 Human NOV	
24	1002	33.5	568	8	ADH42445	ADh42445 Novel hum	
25	1002	33.5	568	8	ADN61775	ADn61775 Human nov	

26	998	33.3	568	7	ADE07994	Ade07994 Novel pro
27	993	33.2	568	5	AAU79946	Aau79946 Human tra
28	993	33.2	568	6	ABG75835	Abg75835 Transport
29	993	33.2	568	8	ADG16993	Adg16993 Human tra
30	985	32.9	616	5	ABU65064	Abu65064 Human NOV
31	985	32.9	616	5	ABU65063	Abu65063 Human NOV
32	985	32.9	616	7	ADK51048	Adk51048 Human NOV
33	985	32.9	616	8	ADH42447	Adh42447 Novel hum
34	985	32.9	616	8	ADN61777	Adn61777 Human nov
35	985	32.9	616	8	ADN61779	Adn61779 Human nov
36	974	32.5	572	8	ADP64791	Adp64791 Rat INDY,
37	966.5	32.3	531	6	ABG75826	Abg75826 Transport
38	965.5	32.2	602	6	ABH82952	Abh82952 Human SLC
39	965.5	32.2	602	6	ABR40101	AbR40101 Human ren
40	965.5	32.2	602	6	ABG75837	Abg75837 Transport
41	959.5	32.0	539	5	AAE21181	Aae21181 Human TRI
42	948.5	31.7	572	6	ABR57024	AbR57024 Mouse TCH
43	948.5	31.7	572	8	ADP64797	Adp64797 Mouse Na+
44	939.5	31.4	551	8	ADR09550	Adr09550 Human pro
45	896	29.9	626	3	AAAB36167	Aab36167 Novel hum
46	895.5	29.9	627	3	AAAB36161	Aab36161 Novel hum
47	891	29.8	626	3	AAAB36164	Aab36164 Novel hum
48	891	29.8	626	3	AAAB36168	Aab36168 Novel hum
49	890.5	29.7	627	3	AAAB36158	Aab36158 Human sec
50	890.5	29.7	627	3	AAAB36162	Aab36162 Novel hum
51	890.5	29.7	627	3	AAAB42213	Aab42213 Human ORF
52	889.5	29.7	627	3	AAAB42213	Aab42213 Human ORF
53	889.5	29.7	627	5	ABB97450	Abb97450 Novel hum
54	886	29.6	626	3	AAAB36165	Aab36165 Novel hum
55	885.5	29.6	627	3	AAAB36159	Aab36159 Novel hum
56	884.5	29.5	519	8	ABM84184	Abm84184 Human dia
57	879.5	29.4	595	5	AAE22910	Aae22910 Human tra
58	879.5	29.4	595	5	AAO21807	Aao21807 Lung-spec
59	879.5	29.4	595	6	ABB82949	Abb82949 Human SLC
60	878.5	29.4	595	7	ADI21045	Adi21045 Novel hum
61	879.5	29.4	595	8	ADO78130	Ado78130 Human SLC
62	875.5	29.2	519	8	ADH22543	Adh22543 Human tra
63	858	28.7	533	8	ABM84303	Abm84303 Human dia
64	852.5	28.5	557	8	ADP29692	Adp29692 Human sec
65	850	28.4	522	5	ABU65065	Abu65065 Human NOV
66	850	28.4	522	7	ADK51054	Adk51054 Human NOV
67	850	28.4	522	8	ADH42449	Adh42449 Novel hum
68	850	28.4	522	8	ADN61781	Adn61781 Human nov
69	819	27.4	520	7	ADM05177	Adm05177 Human pro
70	814.5	27.2	516	5	ABU65066	Abu65066 Human NOV
71	814.5	27.2	516	7	ADK51056	Adk51056 Human NOV
72	814.5	27.2	516	8	ADH42451	Adh42451 Novel hum
73	795	26.6	477	8	ADN61783	Adn61783 Human nov
74	790.5	26.4	551	8	ADP64795	Adp64795 Nematode
75	760	25.4	580	3	AAAB36169	Aab36169 Novel hum
76	759.5	25.4	581	3	AAAB36163	Aab36163 Novel hum
77	755	25.2	580	3	AAAB36166	Aab36166 Novel hum
78	754.5	25.2	581	3	AAAB36160	Aab36160 Novel hum
79	754	25.2	552	6	AAO31005	Aao31005 Human tra
80	632	21.1	552	2	AAW98815	Aaw98815 H. pylori
81	621.5	20.8	540	3	AAAG13103	Aag13103 Arabidops
82	621.5	20.8	558	3	AAAG13102	Aag13102 Arabidops
83	619.5	20.7	540	3	AAAG51055	Aag51055 Arabidops
84	619.5	20.7	557	3	AAAG51054	Aag51054 Arabidops
85	610.5	20.4	510	4	AAAG90000	Aag90000 C glutami
86	605.5	20.2	527	4	AAAB76767	Aab76767 Corynebac
87	603.5	20.2	466	3	AAAG13104	Aag13104 Arabidops
88	601.5	20.1	466	3	AAAG51056	Aag51056 Arabidops
89	598	20.0	543	6	ADA48682	Ada48682 Rice prot
90	598	20.0	543	7	ADC08241	Adc08241 Rice prot
91	595	19.9	432	6	ABM71452	Abm71452 Staphyloc
92	572	19.1	407	5	AAU91115	Aau91115 Human sec
93	572	19.1	407	5	ABG65198	Abg65198 Human alb
94	572	19.1	407	8	ADL78465	Adl78465 Albumin f
95	557	18.6	382	7	ADE09041	Ade09041 Novel pro
96	551.5	18.4	377	5	AAU91090	Aau91090 Human sec
97	551.5	18.4	377	5	ABG65200	Abg65200 Human alb
98	551.5	18.4	377	8	ADL78467	Adl78467 Albumin f

99 542 18.1 502 6 ADA48226 Ada48226 Rice prot

100 478 16.0 304 5 ABB89646 ABB89646 Human pol

ALIGNMENTS

RESULT 1

ABB60315

ID ABB60315 standard; protein; 572 AA.

XX AC ABB60315;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 7737.

XX Drosophila melanogaster polypeptide; cell signalling; insecticide;

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL04418.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

XX Disclosure; SEQ ID NO 7737; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 572 AA;

Query Match 100.0%; Score 2994; DB 4; Length 572;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIEIGEQPPVKCSNFFANHWKGLVFLVPLLCPLVWMLNEGAEFRCMYLLVMAIFW 60

DB 1 MEIEIGEQPPVKCSNFFANHWKGLVFLVPLLCPLVWMLNEGAEFRCMYLLVMAIFW 60

QY 61 VTEALPLYVTSMPIVAFPIGIMSSDQTCRLYFKDTLVMFNGGIMVALAVEYCNLHKRL 120

DB 61 VTEALPLYVTSMPIVAFPIGIMSSDQTCRLYFKDTLVMFNGGIMVALAVEYCNLHKRL 120

QY 121 ALRVIOIVGCSRRRLHFGIMVTMFLSMWISNAACCTAMMCPIIQAVLBELOAQGVCKINH 180

DB 121 ALRVIOIVGCSRRRLHFGIMVTMFLSMWISNAACCTAMMCPIIQAVLBELOAQGVCKINH 180

QY 181 EPOYIVGNGKNNDEPPYPTKITLCYVLGIAYASSLGGCGTIIGTATNLTFFKGIYEAR 240

DB 181 EPOYIVGNGKNNDEPPYPTKITLCYVLGIAYASSLGGCGTIIGTATNLTFFKGIYEAR 240

QY 241 FKNSTEQMDPFTFMFYSVPSMLVYTLITFVFLQWHFMGLWRPKSKBAQEVQREGADVA 300

DB 241 FKNSTEQMDPFTFMFYSVPSMLVYTLITFVFLQWHFMGLWRPKSKBAQEVQREGADVA 300

QY 301 KKVIDORYKDLGPMSEIHEIQVMILFIEMVMYFTRKPGIFLGWADLLNSKOIRNSMPTIF 360

DB 301 KKVIDORYKDLGPMSEIHEIQVMILFIEMVMYFTRKPGIFLGWADLLNSKOIRNSMPTIF 360

QY 361 VVMCFMFLPANYAFRLRYCTRRGGPVPTGTPSLITWKFIOQKVPWGLVFLGGGFALAE 420

DB 361 VVMCFMFLPANYAFRLRYCTRRGGPVPTGTPSLITWKFIOQKVPWGLVFLGGGFALAE 420

QY 421 SKQSGMAKLIQNALIGLKVLPNSVLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAI 480

DB 421 SKQSGMAKLIQNALIGLKVLPNSVLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAI 480

QY 481 ETHPLYLILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIIITLTFVF 540

DB 481 ETHPLYLILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIIITLTFVF 540

QY 541 CQTWGLVVPYNLNSPFEWAQIYAAAALGNKTH 572

DB 541 CQTWGLVVPYNLNSPFEWAQIYAAAALGNKTH 572

RESULT 2

ABB66903

ID ABB66903 standard; protein; 572 AA.

XX AC ABB66903;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 27501.

XX Drosophila melanogaster polypeptide; cell signalling; insecticide;

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL11006.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

XX Disclosure; SEQ ID NO 27501; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 572 AA;

Query Match 100.0%; Score 2994; DB 4; Length 572;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIEIGEQPPVKCSNFFANHWKGLVFLVPLLCPLVWMLNEGAEFRCMYLLVMAIFW 60

DB 1 MEIEIGEQPPVKCSNFFANHWKGLVFLVPLLCPLVWMLNEGAEFRCMYLLVMAIFW 60

QY 61 VTEALPLYVTSMPIVAFPIGIMSSDQTCRLYFKDTLVMFNGGIMVALAVEYCNLHKRL 120

DB 61 VTEALPLYVTSMPIVAFPIGIMSSDQTCRLYFKDTLVMFNGGIMVALAVEYCNLHKRL 120

QY 121 ALRVIOIVGCSRRRLHFGIMVTMFLSMWISNAACCTAMMCPIIQAVLBELOAQGVCKINH 180

DB 121 ALRVIOIVGCSRRRLHFGIMVTMFLSMWISNAACCTAMMCPIIQAVLBELOAQGVCKINH 180

CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX	PA	(UYCO-) UNIV CONNECTICUT.	XX	PI	Rogina B, Reenan RA, Helfand SL;	XX
XX	PI	WPI: 2002-599787/64.	XX	DR	N-PSDB; AEN84439.	XX
XX	DR	Indy polynucleotide, useful for diagnosing or treating body weight	XX	DR	disorders, e.g. obesity, metabolic maintenance disorders, or the symptoms	XX
XX	PT	of aging to extend the life span of an organism.	XX	PT	Claim 14; Page 81-83; 83pp; English.	XX
XX	PS	The present sequence is the protein sequence of the Drosophila	XX	CC	melanogaster INDY protein, which is encoded by the Indy gene (see	XX
XX	CC	ABN84439) involved in increased life span. The protein has similarity to	XX	CC	dicarboxylate transporters such as those from human and rat, and	XX
XX	CC	represents a new class of dicarboxylate transporters that are not	XX	CC	inhibited by phloretin. Identification of the Indy gene resulted from the	XX
XX	CC	observation that particular mutations in the gene caused an increase in	XX	CC	the life span of the fly carrying the mutation. As a result of this	XX
XX	CC	finding, it is now possible to identify and/or isolate Drosophila lines	XX	CC	with longer life spans, as well as to identify agents that contribute to	XX
XX	CC	longer life span. It is also possible to isolate genes involved in, and	XX	CC	which have an effect on longevity, as well as proteins encoded by these	XX
XX	CC	genes. The invention provides Indy polynucleotides, proteins, anti-INDY	XX	CC	antibodies, antagonists that inhibit Indy activity or expression and	XX
XX	CC	agonists that increase Indy activity or expression, and their use in the	XX	CC	diagnosis or treatment of body weight disorders, such as obesity and	XX
XX	CC	metabolic maintenance disorders, or longevity in humans and animals.	XX	CC	Antagonists include at least a portion of the Indy gene sequence, an	XX
XX	CC	antisense oligonucleotide, a ribozyme, a triple helix-forming molecule, a	XX	CC	double-stranded interfering RNA, an anti-Indy antibody, or a mixture of	XX
XX	CC	these. Methods of calorically restricting an organism and of extending	XX	CC	the lifespan of an organism by administering the antagonist are claimed	XX
XX	SQ	Sequence 572 AA;	XX	SQ	Sequence 572 AA;	XX
QY	1	MEIEIGEQQPPVKCSNFFANHWKGLVFLVPLLCPLVMLNEGAEFRCMYLLVMAIFW	QY	1	MEIEIGEQQPPVKCSNFFANHWKGLVFLVPLLCPLVMLNEGAEFRCMYLLVMAIFW	60
DB	1	MEIEIGEQQPPVKCSNFFANHWKGLVFLVPLLCPLVMLNEGAEFRCMYLLVMAIFW	DB	1	MEIEIGEQQPPVKCSNFFANHWKGLVFLVPLLCPLVMLNEGAEFRCMYLLVMAIFW	60
QY	61	VTEALPLYVTSMPDIVAFPPIMGIMSSDQTCRLYFKDTLVMPMGIMVALAVEYCNLHKRL	QY	61	VTEALPLYVTSMPDIVAFPPIMGIMSSDQTCRLYFKDTLVMPMGIMVALAVEYCNLHKRL	120
DB	61	VTEALPLYVTSMPDIVAFPPIMGIMSSDQTCRLYFKDTLVMPMGIMVALAVEYCNLHKRL	DB	61	VTEALPLYVTSMPDIVAFPPIMGIMSSDQTCRLYFKDTLVMPMGIMVALAVEYCNLHKRL	120
QY	121	ALRVIQVGCSPRLHFGIMVTMFLSNWISNAACTAMWCPIIOAVLELOAQGVCKINH	QY	121	ALRVIQVGCSPRLHFGIMVTMFLSNWISNAACTAMWCPIIOAVLELOAQGVCKINH	180
DB	121	ALRVIQVGCSPRLHFGIMVTMFLSNWISNAACTAMWCPIIOAVLELOAQGVCKINH	DB	121	ALRVIQVGCSPRLHFGIMVTMFLSNWISNAACTAMWCPIIOAVLELOAQGVCKINH	180
QY	181	EPQYQIVGKNKKNDEPPYPTKITLCYLGIAVASSLGGCGTIIGTATNLTFRGIYEAR	QY	181	EPQYQIVGKNKKNDEPPYPTKITLCYLGIAVASSLGGCGTIIGTATNLTFRGIYEAR	240
DB	181	EPQYQIVGKNKKNDEPPYPTKITLCYLGIAVASSLGGCGTIIGTATNLTFRGIYEAR	DB	181	EPQYQIVGKNKKNDEPPYPTKITLCYLGIAVASSLGGCGTIIGTATNLTFRGIYEAR	240
QY	241	FKNSTEQMDPPTFMFYSVPMSLVYLLTFVFLQWHFMGLMRPKSKEAQEVQREGADVA	QY	241	FKNSTEQMDPPTFMFYSVPMSLVYLLTFVFLQWHFMGLMRPKSKEAQEVQREGADVA	300
DB	241	FKNSTEQMDPPTFMFYSVPMSLVYLLTFVFLQWHFMGLMRPKSKEAQEVQREGADVA	DB	241	FKNSTEQMDPPTFMFYSVPMSLVYLLTFVFLQWHFMGLMRPKSKEAQEVQREGADVA	300
QY	301	KKVIDQRYKDLGPMSTHEIQWMLFIEMVVMYFTRKPGIFLGWADLLNSKDTRNSMPTIF	QY	301	KKVIDQRYKDLGPMSTHEIQWMLFIEMVVMYFTRKPGIFLGWADLLNSKDTRNSMPTIF	360
DB	301	KKVIDQRYKDLGPMSTHEIQWMLFIEMVVMYFTRKPGIFLGWADLLNSKDTRNSMPTIF	DB	301	KKVIDQRYKDLGPMSTHEIQWMLFIEMVVMYFTRKPGIFLGWADLLNSKDTRNSMPTIF	360
QY	361	VVMCFMLPANYAFRLRYCTRRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAEG	QY	361	VVMCFMLPANYAFRLRYCTRRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAEG	420
DB	361	VVMCFMLPANYAFRLRYCTRRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAEG	DB	361	VVMCFMLPANYAFRLRYCTRRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAEG	420
QY	421	SKQSGMAKLIGNALIGLKVLPNSVLLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAI	QY	421	SKQSGMAKLIGNALIGLKVLPNSVLLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAI	480

RESULT 3  
ABB79611  
ID ABB79611 standard; protein; 572 AA.  
XX ABB79611;  
AC ABB79611;  
DT 21-OCT-2002 (first entry)  
XX Drosophila INDY protein dicarboxylate transporter.  
DE Ind; dicarboxylate transporter; life span; longevity; obesity;  
KW anorectic; caloric restriction; transgenic animal; gene therapy.  
XX Drosophila melanogaster.  
OS WO200259310-A2.  
XX 01-AUG-2002.  
PD 12-DEC-2001; 2001WO-US0481130.  
XX 12-DEC-2000; 2000US-0255013P.

Db 421 SKOSGMAKLIGNALIGLKVLPSVLLVWLVAVFLTAFSSNVAIANIIPVLAEMSLAI 480  
QY 481 EIHPYLYLILPAGLACSMFHLPVSTPNALVAGYANIRTKDMAIAGIGPTIITITLFFVF 540  
Db 481 EIHPYLYLILPAGLACSMFHLPVSTPNALVAGYANIRTKDMAIAGIGPTIITITLFFVF 540  
QY 541 CQTWGLVVPNLNSFPPEWAIYAAAALGNKTH 572  
Db 541 CQTWGLVVPNLNSFPPEWAIYAAAALGNKTH 572

RESULT 4  
ADP64789  
ID ADP64789 standard; protein; 572 AA.  
XX AC  
XX ADP64789;  
XX DT 26-AUG-2004 (first entry)  
XX DE Drosophila INDY, an Na<sup>+</sup>-coupled citrate transporter protein.  
XX KW sodium-coupled citrate transporter; transmembrane citrate transporter;  
XX KW lifespan; weight reduction; weight gain prevention; blood cholesterol;  
XX KW triglyceride; low density lipopolysaccharide; glucose; obesity;  
XX KW hyperlipidemia; hypercholesterolemia; INDY protein.  
XX OS Drosophila melanogaster.  
XX PN WO2004048925-A2.  
XX PD 10-JUN-2004.  
XX PF 20-NOV-2003; 2003WO-US037054.  
XX PR 22-NOV-2002; 2002US-0428469P.  
XX PR 01-APR-2003; 2003US-0459441P.  
XX PA (MEDI-) MEDICAL COLLEGE GEORGIA RES INST.  
XX PA (GANA/) GANAPATHY V.  
XX PA (INOU/) INOUE K.  
XX PA (FEIY/) FEI Y.  
XX PI Ganapathy V, Inoue K, Fei Y;  
XX WPI; 2004-460797/43.  
XX N-PSDB; ADP64788.  
XX New isolated polynucleotide encoding a Na<sup>+</sup>-coupled citrate transporter  
PT (NACT) polypeptide, useful as a drug target for the treatment of obesity,  
PT hyperlipidemia, and hypercholesterolemia.  
XX Claim 37; SEQ ID NO 2; 186pp; English.  
XX The invention relates to novel Na<sup>+</sup>-coupled citrate transporter proteins  
CC and their encoding genes. Inhibitors of transmembrane citrate  
CC transporters are useful for extending the lifespan, reducing weight,  
CC preventing weight gain or lowering blood cholesterol, triglyceride, LDL  
CC or glucose levels in a subject. The NACT polypeptide is useful as a drug  
CC target for the treatment of obesity, hyperlipidemia, and  
CC hypercholesterolemia. This sequence corresponds to the Drosophila INDY  
CC protein.  
XX SQ Sequence 572 AA;

Query Match 100.0%; Score 2994; DB 8; Length 572;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIEIGEOPQPVKCSNFFANHWKGLVPLVPLLCIPVMLLNEGAEFRCMYLLVMAIFW 60  
Db 1 MEIEIGEOPQPVKCSNFFANHWKGLVPLVPLLCIPVMLLNEGAEFRCMYLLVMAIFW 60  
QY 61 VTEALPLYVTSMTPIVAFPPINGIMSSDQTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRL 120

Db 61 VTEALPLYVTSMTPIVAFPPINGIMSSDQTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRL 120  
QY 121 ALRVIQIVGCSPPRLHFGILINVTWFLSMWISNAACATAMCPIIOAVLELOAQGVCKINH 180  
Db 121 ALRVIQIVGCSPPRLHFGILINVTWFLSMWISNAACATAMCPIIOAVLELOAQGVCKINH 180  
QY 181 EPQYQIVGNGKKNNEDEPPYPTKITLCYVLGIAYASSLGCGGTIIGTATNLTFKGIYEAR 240  
Db 181 EPQYQIVGNGKKNNEDEPPYPTKITLCYVLGIAYASSLGCGGTIIGTATNLTFKGIYEAR 240  
QY 241 FKNSTEQMDPFTFMEYSVPSMLVYLLTFVFLQWHFMGLWRPKSKEAOEVORREGADVA 300  
Db 241 FKNSTEQMDPFTFMEYSVPSMLVYLLTFVFLQWHFMGLWRPKSKEAOEVORREGADVA 300  
QY 301 KKVIDORYKDLGPMISHEIQWMLFIEMVVMVFTKPGIFLGWADLLNSKDINSMPTIF 360  
Db 301 KKVIDORYKDLGPMISHEIQWMLFIEMVVMVFTKPGIFLGWADLLNSKDINSMPTIF 360  
QY 361 VVVMCFMPLPANYAFRLRYCTRRGPPVPTGPTPSLITWKFITQTKVPWGLVFLGGGFALAEG 420  
Db 361 VVVMCFMPLPANYAFRLRYCTRRGPPVPTGPTPSLITWKFITQTKVPWGLVFLGGGFALAEG 420  
QY 421 SKOSGMAKLIGNALIGLKVLPSVLLVWLVAVFLTAFSSNVAIANIIPVLAEMSLAI 480  
Db 421 SKOSGMAKLIGNALIGLKVLPSVLLVWLVAVFLTAFSSNVAIANIIPVLAEMSLAI 480  
QY 481 EIHPYLYLILPAGLACSMFHLPVSTPNALVAGYANIRTKDMAIAGIGPTIITITLFFVF 540  
Db 481 EIHPYLYLILPAGLACSMFHLPVSTPNALVAGYANIRTKDMAIAGIGPTIITITLFFVF 540  
QY 541 CQTWGLVVPNLNSFPPEWAIYAAAALGNKTH 572  
Db 541 CQTWGLVVPNLNSFPPEWAIYAAAALGNKTH 572

RESULT 5  
ABR40099  
ID ABR40099 standard; protein; 572 AA.  
XX AC ABR40099;  
XX DT 24-JUL-2003 (first entry)  
XX DE dIndyaa sequence.  
XX KW Anti-diabetic; anorectic; sodium dicarboxylate cotransporter;  
XX KW sodium tricarboxylate cotransporter; HepNaDc; hepatocyte; HepG2;  
XX KW diabetes; obesity; lipid metabolism; aging; dIndyaa.  
XX OS Unidentified.  
XX PN WO2003029465-A1.  
XX PD 10-APR-2003.  
XX PF 27-SEP-2002; 2002WO-JP010038.  
XX PR 28-SEP-2001; 2001JP-00299433.  
XX PR 28-AUG-2002; 2002JP-00249016.  
XX PA (SAKA ) OTSUKA PHARM CO LTD.  
XX Kanemoto N, Omori Y, Sugano S, Obuchi Y;  
XX WPI; 2003-354728/33.  
XX Sodium di- or tricarboxylate cotransporter gene (HepNaDC) for treatment  
PT and prevention of diabetes, obesity, for improving lipid metabolism and  
PT aging.  
XX Example 1; Fig 1; 75pp; Japanese.  
XX

CC The present invention relates to human sodium di- or tricarboxylate  
CC cotransporter protein (HepNadC; AB40097). HepNadC gene is expressed in  
CC human hepatocyte cancer-origin cells HepG2. Compounds which control the  
CC expression of the HepNadC gene and activity of the HepNadC polypeptide  
CC may be useful for treating and preventing diabetes, obesity, for  
CC improving lipid metabolism and aging. The present sequence is the dindaa  
CC protein, which was used in a sequence alignment with the HepNadC sequence  
XX  
SQ

Query Match 99.7%; Score 2986; DB 6; Length 572;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 571; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEIEIGEQQPPVKCSNFANHWKGLVFLVLLCLPVMMLNEGAFRCMYLLLVNAIFW 60  
DB 1 MEIEIGEQQPPVKCSNFANHWKGLVFLVLLCLPVMMLNEGAFRCMYLLLVNAIFW 60  
QY 61 VTEALPLYVTSMIPVAFPIGMSDQTCRLYFKDTLVNFMGGIMVALAVYCNLHKEL 120  
DB 61 VTEALPLYVTSMIPVAFPIGMSDQTCRLYFKDTLVNFMGGIMVALAVYCNLHKEL 120  
QY 121 ALRVQIVGCSRRRLHFGIMVTMFLSMISNAACTAMMCPPIQAVLELQAGVCKINH 180  
DB 121 ALRVQIVGCSRRRLHFGIMVTMFLSMISNAACTAMMCPPIQAVLELQAGVCKINH 180  
QY 181 EPQYQIVGKNKNNDEPPYPTKITLCYVIGIAYASSLGGCGTIIGTATNLFKGIYEAR 240  
DB 181 EPQYQIVGKNKNNDEPPYPTKITLCYVIGIAYASSLGGCGTIIGTATNLFKGIYEAR 240  
QY 241 FKNSTEQMDPFTFMFVSVMVLYLLTFVFLQWHFMGLWRPKSKEAQVQRGREGADVA 300  
DB 241 FKNSTEQMDPFTFMFVSVMVLYLLTFVFLQWHFMGLWRPKSKEAQVQRGREGADVA 300  
QY 301 KKVVDQRYKDLGPMSTHEIQVMILFIEMVVMYFTRKPGIFLGWADLLNSKDIRNSMPTIF 360  
DB 301 KKVVDQRYKDLGPMSTHEIQVMILFIEMVVMYFTRKPGIFLGWADLLNSKDIRNSMPTIF 360  
QY 361 VVMCFMFLPANYAFRYCTRRGGPVTGTPSLITWKFTQTKVPWGLVFLGGGFALAE 420  
DB 361 VVMCFMFLPANYAFRYCTRRGGPVTGTPSLITWKFTQTKVPWGLVFLGGGFALAE 420  
QY 421 SKQSGMAKLIGNALIGLKVLPNSVLLVILVAVELTAPSSNVAIANIIPVLAEMSLAI 480  
DB 421 SKQSGMAKLIGNALIGLKVLPNSVLLVILVAVELTAPSSNVAIANIIPVLAEMSLAI 480  
QY 481 EIHPLYLILPAGLACSMAPHLVPSTPPNALVAGYANIRTKDMAIAGIGPTIITITLTVF 540  
DB 481 EIHPLYLILPAGLACSMAPHLVPSTPPNALVAGYANIRTKDMAIAGIGPTIITITLTVF 540  
QY 541 CQTWGLVVPNLNSPPEWAIYAAAALGNKTH 572  
DB 541 CQTWGLVVPNLNSPPEWAIYAAAALGNKTH 572

RESULT 6  
ABB61068  
ID ABB61068 standard; protein; 908 AA.  
XX  
AC ABB61068;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 9996.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.  
XX 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX (PEKE) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX MPI; 2001-656860/75.  
DR N-PSDB; ABL05171.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
PS Disclosure; SEQ ID NO 9996; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB116176-ABL30511), expressed DNA  
CC sequences (AB101840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 908 AA;  
Query Match 44.3%; Score 1327.5; DB 4; Length 908;  
Best Local Similarity 46.8%; Pred. No. 5.6e-136;  
Matches 253; Conservative 99; Mismatches 135; Indels 54; Gaps 7;  
QY 22 HWKGLVFLVLLCLPVMMLNEGAFRCMYLLLVNAIFWTEALPLYVTSMIPVAFPIM 81  
DB 418 HWKGLVFLVLLCLPVMMLNEGAFRCMYLLLVNAIFWTEALPLYVTSMIPVAFPIM 81  
QY 82 GIM--SSDQTCRLYFKDTLVNFMGGIMVALAVYCNLHKRLALRVQIVGCSRRRLHGL 139  
DB 451 DILESTSEKVSFYFSDTVNFMFGGLTALAEVSNLHQRALNTLIVGCSRRRLHGL 510  
QY 140 IMVTMFLSMISNAACTAMMCPPIQAVLELQAGVCKINHHEPOYQIVGKNKNNDEPP 199  
DB 511 VMTCTSLMISNAACTAMMCPPIQAVLELQAGVCKINHHEPOYQIVGKNKNNDEPP 564  
QY 200 YPTKITLCYVIGIAYASSLGGCGTIIGTATNLFKGIYEARFPKSTEQMDPFTFMFVSVP 259  
DB 565 HPSTISMAFYGIAYSSSISGCGTLIGTNTLVKGLYDTRFPNSDEKIDPFIEMAYSVP 624  
QY 260 -SMLVYLLTFVFLQWHFMGLWRPKSKEAQVQRGREGADVAKKVIDORYKDLGPMSTHE 318  
DB 625 FVILVILFLTYFLSLOVTHMGLFRPNSKIGQVKKGAESQDVVKDVIKORKAELGPMSCHE 684  
QY 319 IQVMILFIEMVVMYFTRKPGIFLGWADLLNSKDIRNSMPTIFVVMCFMFLPANYAFRLYC 378  
DB 685 IQVGLLVILVILVAVELTAPSSNVAIANIIPVLAEMSLAIEHPLYLILPAGLACSM 744  
QY 379 TRRGVPVPTGTPSLITWKFTQTKVPWGLVFLGGGFALAEQSGSKQMAKIGNALIGLK 438  
DB 745 CGK-APFPGQTLDACLSW-----GPALESGRSVSGWAKMGLGESLAFAG 786  
QY 439 VLPNSVLLVILVAVELTAPSSNVAIANIIPVLAEMSLAIEHPLYLILPAGLACSM 498  
DB 787 EMHSLVISMCIILSLFCTAFASNAICNLIPFSEMAIAIEVHPMKLTTPPALACSLA 846  
QY 499 FHLVSTPPNALVAGYANIRTKDMAIAGIGPTIITITLTVFVCCQTVGLVVPNLNSPPEW 558  
DB 847 FHLVSTPPNALVAGYANIRTKDMAIAGIGPTIITITLTVFVCCQTVGLVVPNLNSPPEW 906  
QY 559 A 559

QY	16	LVVFLVPLLCPLVMLNEGAFRCMYLLVMAIFWVTEALPLYVVTSMPIVIAFPPIGIMS	85
Db	15	LIVCLPIPLPLPLIVQTKEAYCAYSIIIMALLWCTEALPLAVTALFPIVLPLMGIMD	74
QY	86	SDQTCRLYFKDTLVNFMGGIMVAVENLHKLRLAVIQIVGCSPPRLHFLMTWTF	145
Db	75	ASEVCIEYFKDTNILFVGGLMVAIAVEHNLHKLRLAVIQIVGCSPPRLHFLMTWTF	134
QY	146	LSMWISNDACTAMMCPILIOAVLELQA--QGCKLNHPEPOYIV-----GGNK	191
Db	135	LSMWISNTATTAMMVPFGHVALEQGGKKDVEGNNPTFELQECQKEVTKLDNGOP	194
QY	192	KNNEDEP-----PYPTKITLTCYLGIAIYASSLGCGGTIIGTATNLTFKGIYEARPK	242
Db	195	VSAPEPRTKQTEHHRFSQGLSLC-----ICVSASIGGIATLTGTPNLVLQGVNSLFP	250
QY	243	NETEQMDPTFMFYSVPMSLVYTLITFVLOHFMGLWRPKSKEAQEVQREGADVAKK	302
Db	251	QNGNVNFAFWGFAFPTMIILLLAWLWQLVFLGVNFRKNFGFGGEGEERKQA--AFQ	308
QY	303	VIDQYKDLGPMISHEIQVMILFIWVWYFTRKPGIFLGWADLLNSKDINSMT-----	358
Db	309	VIKTQRLGLGPMISHEIQVMILFIWVWYFTRKPGIFLGWADLLNSKDINSMT-----	368
QY	359	-IFVVMCFMPLPANYAFRLYCTRGRGPVPTGTPSLITWKFTQTKVPMGLVFLGGFAL	417
Db	369	ALFISLVMPFIIPSKIPGLMQDPKPKLKA--PPAILTWKTVDNKPMPNIVILGGGFAL	426
QY	418	AGSKQSGMAKLIGNALIGLVPLNSVLLVILVAVFLTAFFSSNVAIANIIPVLAEMS	477
Db	427	AKGSEQSGLEWLDKLTPLQHPSPSATVILICLLAIETECTSNVATTLFLPILASNA	486
QY	478	LATEIHPLYLILPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITL	537
Db	487	QAICLHPLVMPCTLAASLAFMLPVATPPNAIVFSFGGLKVSMDMARAGFLNIIIGLAI	546
QY	538	FVFCQWGLVVPVNLNSPPEWAQIYAAAAALGNKTH	572
Db	547	TLISNSWSIPIF-KLDTFSPWAHSNTSQCLLNPSN	580
RESULT 8			
ID	ADP64799	standard; protein; 581 AA.	
XX	AC	ADP64799;	
XX	XX	26-AUG-2004 (first entry)	
DE	XX	Zebrafish Na+-coupled citrate transporter protein.	
KW	KW	sodium-coupled citrate transporter; transmembrane citrate transporter;	
KW	KW	lifespans; weight reduction; weight gain prevention; blood cholesterol;	
KW	KW	triglyceride; low density lipoprotein; glucose; obesity;	
XX	XX	hyperlipidemia; hypercholesterolemia; INDY protein.	
OS	XX	Danio rerio.	
XX	XX	WO2004048925-A2.	
PD	XX	10-JUN-2004.	
PF	XX	20-NOV-2003; 2003WO-US037054.	
XX	XX	22-NOV-2002; 2002US-0428469P.	
PR	XX	01-APR-2003; 2003US-0459441P.	
XX	XX	(MEDI-) MEDICAL COLLEGE GEORGIA RES INST.	
PA	XX	(GANA/) GANAPATHY V.	
PA	XX	(INOU/) INOUE K.	
XX	XX	(FEIY/) FEI Y.	
PI	XX	Ganapathy V, Inoue K, Fei Y;	

Db 907 A 907

RESULT 7

AAE38766

ID AAE38766 standard; protein; 587 AA.

XX AC AAE38766;

XX XX 18-DEC-2003 (first entry)

XX XX Rat 69624 protein sodium sulphate symporter domain.

XX XX Rat; 69624; transporter protein; neurological disorder; therapy;

KW arteriosclerosis; cardiac hypertrophy; ischaemia reperfusion injury;

KW metastatic disorder; haematopoietic neoplastic disorder; leukaemia;

KW arthritis; multiple sclerosis; encephalomyelitis; myasthenia gravis;

KW carcinoma; cell proliferation; autoimmune disorder; diabetes mellitus;

KW renal disorder; colon; hepatic disease; hypocalcaemia; calcium stone;

KW mental retardation; Canavan disease; differentiative disorder; sarcoma;

KW systemic lupus erythematosus; cardiovascular disorder; arteriosclerosis;

KW atrial fibrillation; forensic identification; pain.

XX XX Rattus norvegicus.

OS XX US2002193582-A1.

PN XX 19-DEC-2002.

PD XX 17-JUN-2002; 2002US-00173519.

PF XX 18-JUN-2001; 2001US-0298970P.

PR XX (MILL-) MILLENNIUM PHARM INC.

XX PA Curtis RAJ;

XX PI WPI; 2003-644620/61.

XX DR Novel 69624 polypeptide, a human transporter family member, useful for

XX PT treating disorders e.g. hypocalcaemia, such as hypocalcaemia, formation

XX PT of calcium stones, mental retardation abnormal body sulfate homeostasis.

XX PS Disclosure; Page 49-50; Opp; English.

XX XX The invention relates to 69624 polypeptide, a human transporter family

CC member and its corresponding nucleic acid. 69624 protein is useful for

CC developing novel diagnostic and therapeutic agents for 69624-mediated or

CC related disorders. 69624 protein act as therapeutic or diagnostic agents

CC for renal, neurological, colon or hepatic disorders. It act as diagnostic

CC targets and therapeutic agents for treating disorders such as

CC hypocalcaemia, formation of calcium stones, mental retardation (Canavan

CC disease) or abnormal body sulfate homeostasis. 69624 protein may act as

CC diagnostic targets and therapeutic agents for controlling cellular

CC proliferative and/or differentiative disorders such as carcinoma,

CC sarcoma, metastatic disorder or haematopoietic neoplastic disorders e.g.,

CC leukaemia, immune disorders such as autoimmune disorders (diabetes

CC mellitus, arthritis), multiple sclerosis, encephalomyelitis, myasthenia

CC gravis, systemic lupus erythematosus, cardiovascular disorders such as

CC arteriosclerosis, atherosclerosis, ischaemia reperfusion injury, cardiac

CC hypertrophy, atrial fibrillation etc; and disorders involving abnormal or

CC excessive pain. 69624 sequence is useful as pharmacodynamic marker and is

CC also used in forensic identification of a biological sample. The present

CC sequence is rat sodium-dependent dicarboxylate/sulphate transporter (NaDC

CC -1) protein

XX XX Sequence 587 AA;

XX XX Query Match

XX XX Best Local Similarity 34.8%; Score 1042.5; DB 7; Length 587;

XX XX Matches 219; Conservative 129; Mismatches 190; Indels 37; Gaps 8;



Db 306 EQQAAVCVITQTEHRLGPMTEAEKASISILFVLVLLWFTTRPGFGLGKGNLAFNPAKGE 365  
 QY 352 --TRNSMPTIFVVMCMFLPANTAFRLYCTRRGGPV--PTGPTFSLITWKFIOTKVPWGL 407  
 Db 366 SMVSDGTVAIFIGIIMFISKPGLTQDPENFGKLKAPLG---LLDWMKTVNQKMPWNI 421  
 QY 408 VFLGGGFALAEKSKQSGMAKLIGNALIGLVKVPNSVLLVILVAVLFTAFSSNVAIAN 467  
 Db 422 VLLGGGYALAKGSGSLSEWLNKLTPLQSVPAFAIAIILSLVATFTECTSNVATT 481  
 QY 468 IIIIVLAEMSLATIHPLYLILPAGLACSMAPHLPVSTPPNALVAGVANRTKDMAIAGI 527  
 Db 482 IFLPILASMAQAICLHPLYVMLPCTLATSLAFMLPVATPPNAIVFSGDLKVLDMARAGF 541  
 QY 528 GPTIITITITLFFVFCQWGLVVPNLSFPEWAQIYAAA---ALGNKT 571  
 Db 542 LLNIIGVLIILALAINSGIPLF-SLHSFSPSAQSNNTTAQCLPSLANTT 588

RESULT 10  
 ABB82951  
 ID ABB82951 standard; protein; 592 AA.  
 XX AC ABB82951;  
 XX DT 14-APR-2003 (first entry)  
 XX DE Human SLC13A related protein (GenBank Identifier No. GI#4506979).  
 XX KW SLC13A; p53; sodium-sulfate cotransporter 2; cytostatic; cancer;  
 XX KW transmembrane protein; human.  
 XX OS Homo sapiens.  
 XX PN WO200298468-A1.  
 XX PD 12-DEC-2002.  
 XX PF 03-JUN-2002; 2002MO-US017460.  
 XX PR 05-JUN-2001; 2001US-0296076P.  
 XX PR 10-OCT-2001; 2001US-0328605P.  
 XX PR 15-FEB-2002; 2002US-0357253P.  
 XX PR 01-MAR-2002; 2002US-0361196P.  
 XX PA (EXEL-) EXELIXIS INC.  
 XX PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;  
 XX PI Lioubin MN;  
 XX DR WPI; 2003-167297/16.  
 XX DR N-PSDB; AB224204.  
 XX PT Identifying candidate p53 pathway modulating agent for diagnosing or  
 XX PT treating cancer comprises detecting test agent-biased activity of an  
 XX PT assay system comprising purified Sodium sulfate cotransporter 2 (SLC13A)  
 XX PT polypeptide.  
 XX PS Claim 13; Page 62-64; 69pp; English.  
 XX CC The invention relates to identifying a candidate p53 pathway modulating  
 XX CC agent. The method involves assaying purified Sodium-sulfate cotransporter  
 XX CC 2 (SLC13A) polypeptide or nucleic acid or a functionally active fragment  
 XX CC or derivative with a test agent. The methods are useful for identifying a  
 XX CC candidate p53 pathway modulating agent, modulating a p53 pathway of a  
 XX CC cell, or a mammalian cell and for diagnosing a disease in a patient. The  
 XX CC disease is breast, colon, lung or ovary cancer having greater than 25%  
 XX CC expression level. The method is useful for manufacturing a medicament for  
 XX CC diagnosing or treating breast, colon, lung or ovary cancer. Sequences  
 XX CC ABB82949-952 represent polypeptide sequences related to the human SLC13A  
 XX CC protein  
 XX SQ Sequence 592 AA;

Query Match 34.3%; Score 1026; DB 6; Length 592;  
 Best Local Similarity 37.2%; Pred. No. 4.e-103; Indels 56; Gaps 10;  
 Matches 219; Conservative 124; Mismatches 189;  
 QY 26 LVVFLVPLLCPLPMLLNEGABFRCMYLLLVMAIFWVTEALPLYVTSMIPVIAFPIMGIMS 85  
 Db 15 LIVFFVILLPLPLTPVPSKEAYCAIILMALFWCTEALPLAVTALPLILFPMKVID 74  
 QY 86 SDQTRLVFKDTLVNFMGGIIVWALAVEYCNLHKRLALRVIQIVGSCSPRLHGLIMVTFM 145  
 Db 75 ASEVAEYELKDSNLLFFGGLLVIAIAVEHNLKRLALRVLLIVGVRRPAPLIILGFLMVTAF 134  
 QY 146 LGMWISNAACTAMMCPILQAVLEELQAQGVCKINHEPOVQIVGNGKKNEDDEPPYTK-- 203  
 Db 135 LSWISNATATSMWVPIAHAVLDQLHSSQASSNVSE-----GSNNPTFELQPSPOKEV 188  
 QY 204 -----ITLCYYIGIAYASSLGGCGTIIIGTATNITFKG 235  
 Db 189 TKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCMSLCVCYSASIGGIATLTGTAPNLVQG 248  
 QY 236 IYEAFKNSQTEQMDPTTFMFYSPSMLVYTLTLTFVLOWHFMGL-WRPKSKEAQEVQGR 294  
 Db 249 QINSUFPQNGVNVNFASFWSFAPPTWVILLALLLAWLQILFLGFNFRKNGIGERKQ--- 305  
 QY 295 EGADVAKVIDORYKDLGPMSTHEIQVMILFIEMVVMYFTRKPGIFLGWADLL--NSKD- 351  
 Db 306 EQQAAVCVITQTEHRLGPMTEAEKASISILFVLVLLWFTTRPGFGLGKGNLAFNPAKGE 365  
 QY 352 --IRNSMPTIFVVMCMFLPANTAFRLYCTRRGGPV--PTGPTFSLITWKFIOTKVPWGL 407  
 Db 366 SMVSDGTVAIFIGIIMFISKPGLTQDPENFGKLKAPLG---LLDWMKTVNQKMPWNI 421  
 QY 408 VFLGGGFALAEKSKQSGMAKLIGNALIGLVKVPNSVLLVILVAVLFTAFSSNVAIAN 467  
 Db 422 VLLGGGYALAKGSGSLSEWLNKLTPLQSVPAFAIAIILSLVATFTECTSNVATT 481  
 QY 468 IIIIVLAEMSLATIHPLYLILPAGLACSMAPHLPVSTPPNALVAGVANRTKDMAIAGI 527  
 Db 482 IFLPILASMAQAICLHPLYVMLPCTLATSLAFMLPVATPPNAIVFSGDLKVLDMARAGF 541  
 QY 528 GPTIITITITLFFVFCQWGLVVPNLSFPEWAQIYAAA---ALGNKT 571  
 Db 542 LLNIIGVLIILALAINSGIPLF-SLHSFSPSAQSNNTTAQCLPSLANTT 588

RESULT 11  
 ABR40100  
 ID ABR40100 standard; protein; 592 AA.  
 XX AC ABR40100;  
 XX DT 24-JUL-2003 (first entry)  
 XX DE Human renal sodium dicarboxylate co-transporter NADCL.  
 XX KW Anti-diabetic; anorectic; sodium dicarboxylate cotransporter; human;  
 XX KW sodium tricarboxylate cotransporter; HepNadC; hepatocyte; HepG2;  
 XX KW diabetes; obesity; lipid metabolism; aging; NADCL.  
 XX OS Homo sapiens.  
 XX PN WO2003029465-A1.  
 XX PD 10-APR-2003.  
 XX PF 27-SEP-2002; 2002MO-JP010038.  
 XX PR 28-SEP-2001; 2001JP-00299433.  
 XX PR 28-AUG-2002; 2002JP-00249016.  
 XX PA (SAKA ) OTSUKA PHARM CO LTD.  
 XX SQ

PI Kanemoto N, Omori Y, Sugano S, Obuchi Y;  
DR WPI; 2003-354728/33.  
XX  
XX Sodium di- or tri-carboxylate cotransporter gene (HepNADC) for treatment  
PT and prevention of diabetes, obesity, for improving lipid metabolism and  
PT aging.  
XX  
XX Example 1; Fig 1; 75pp; Japanese.  
XX  
XX The present invention relates to human sodium di- or tri-carboxylate  
CC cotransporter protein (HepNADC; ABR40097). HepNADC gene is expressed in  
CC human hepatocyte cancer-origin cells Hep32. Compounds which control the  
CC expression of the HepNADC gene and activity of the HepNADC polypeptide  
CC may be useful for treating and preventing diabetes, obesity, for  
CC improving lipid metabolism and aging. The present sequence is the NADC1  
CC protein (GenBank U26209), which was used in a sequence alignment with the  
CC HepNADC sequence  
XX  
XX Sequence 592 AA;  
SQ  
Query Match 34.3%; Score 1026; DB 6; Length 592;  
Best Local Similarity 37.2%; Pred. No. 4.4e-103;  
Matches 219; Conservative 124; Mismatches 189; Indels 56; Gaps 10;  
QY 26 LVVFLVPLCLPVMMLNEGAEFCMVLVMAIFWVTEALPLVYTSMTPIVAFPPINGIMS 85  
DB 15 LIVFVPIILLPLILVPSKEAYCAVAILMALFWCTEALPLAVTALPFLILFPMGIVD 74  
QY 86 SDQTCRLYFKDPLVFMFGIMVALAVEYCNLHKRLARVIOVGCSPRLHFLGIMVTF 145  
DB 75 ASEVAVEYKDSNLLFFGGLLVAFAVEHNLKRLARVLIIVGRPAPLPLGFMVTF 134  
QY 146 LSMWISNAACTAMMCPPIQAVLEELQAGVCKINHEPQYIVGNGKNKNEDEPPYTK-- 203  
DB 135 LSMWISNATSAVMVPIAHAVLDQLHSSQASNVVE-----GNNPTFELQEPSQKEV 188  
QY 204 -----ITLCYVILGIAYASSLGGCGTIGPATNLTPKG 235  
DB 189 TKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCMSLCVCYSASIGGIATLTGTAPNLVLQ 248  
QY 236 IYEAEPKNSQNDPPTFMFYVSPMLVYLLTFVFLQHFHGL-WRPKSKAEQVQRGR 294  
DB 249 QINSFLPQNGVNVNFASFSPFTMTVILLLAWLQTLFLGFNFRKNFGIGERQV--- 305  
QY 295 EGADVAKKVIDQRYKDLGPMSTHEIQVMTLFTFMVVMYFTRKPGIFLGWADLL--NSKD- 351  
DB 306 EQQQAAYCVIQTERRLLGPMTPAEKASILFVLVLLWFTREPFGFLGWNLAFFNAKGE 365  
QY 352 --IRNSMPTIFVVMCMFLPANYAFRLYCTRRGGPV--PTGPTPSLITWKFTQTQVPMGL 407  
DB 366 SMVSDGTVAIFIGIIMFIIPSKFPGITQDPENPGKAPLG---LLDKVTNQKMPNI 421  
QY 408 VPLGGGFALAGSQSGMAKIGNALIGLKVLPNSVLLVVLVAVFLTAPSSNVAIAN 467  
DB 422 VLLGGVALAGKSGRSGLSEGNKGLTPQSQVPAPATAIILSLVAVTFECTSNVATT 481  
QY 468 IIPVLAESLAIEIHPYLLIPAGLACSMAPHLVSTPPNALVAGYANIRTKMAIAGI 527  
DB 482 IFLPLASMAQAICLHPYVLMPLPCTLATSLAPMLPVATPPNNAIVPSFGDLKVLDMARAGF 541  
QY 528 GPTIITITLTFVFCOTGLVGVVNLNSPPEWQAIIYAAA-----ALGNKT 571  
DB 542 LNIIGVLIILALINSWGIPLF-SLHSFSPSAQSNTTAQCLPSLANTT 588  
RESULT 12  
ABR57025  
ID ABR57025 standard; protein; 592 AA.  
XX  
AC ABR57025;  
XX  
DT 05-AUG-2003 (first entry)

XX Human NADC-1 amino acid sequence.  
DE  
XX  
KW Human; TCH169; dicarboxylate transport; hepatotropic; cytostatic;  
KW nephrotropic; vasotropic; antidiabetic; liver disease; hepatitis;  
KW hepatic sclerosis; alcohol-related liver disease; prostate disease;  
KW prostatitis; prostatic hypertrophy; spleen disease; spleen hyperactivity;  
KW kidney disease; nephritis; kidney failure; nephritis; dropsy; diabetes;  
KW diabetes-associated renal disease; metabolic disease; hyperlipaemia;  
KW circulatory disease; arteriosclerosis; cancer; NADC-1.  
OS  
XX Homo sapiens.  
XX  
XX WO2003025168-A1.  
PN  
XX 27-MAR-2003.  
PD  
XX 13-SEP-2002; 2002WO-JP009444.  
PF  
XX 17-SEP-2001; 2001JP-00281992.  
PR  
XX 02-OCT-2001; 2001JP-00306873.  
PR  
XX 16-APR-2002; 2002JP-00113279.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX  
XX Nakanishi A, Uno Y, Sagiya Y;  
PI  
XX WPI; 2003-313352/30.  
DR  
XX  
XX Protein TCH169 with dicarboxylate transport activity for treatment and  
PT diagnosis of diseases including liver disease, cancer and circulatory  
PT disorders.  
XX  
XX Example 1; Fig 1-2; 132pp; Japanese.  
PS  
XX  
XX The present invention describes protein TCH169 and its salts having  
CC dicarboxylate transport activity. TCH169 has hepatotropic, cytostatic,  
CC nephrotropic, vasotropic and antidiabetic activities. The TCH169 protein  
CC and polynucleotide can be used in the treatment, prevention and diagnosis  
CC of liver disease (such as hepatitis, hepatic sclerosis and alcohol-  
CC related liver disease); prostate disease (such as prostatitis and  
CC prostatic hypertrophy); spleen disease (such as spleen hyperactivity);  
CC kidney disease (such as nephritis, kidney failure, nephritis, dropsy and  
CC diabetes-associated renal disease); metabolic disease (such as diabetes);  
CC circulatory disease (such as hyperlipaemia and arteriosclerosis); and  
CC cancer (such as non-small cell lung cancer, liver cancer, renal cancer,  
CC ovarian cancer, prostate cancer, stomach cancer, pancreatic cancer,  
CC breast cancer, colon cancer, bladder cancer and womb cancer). The present  
CC sequence represents a human NADC-1 amino acid sequence which is given in  
CC comparison with the human TCH169 amino acid sequence in an example from  
CC the present invention  
XX  
XX Sequence 592 AA;  
SQ  
Query Match 34.3%; Score 1026; DB 6; Length 592;  
Best Local Similarity 37.2%; Pred. No. 4.4e-103;  
Matches 219; Conservative 124; Mismatches 189; Indels 56; Gaps 10;  
QY 26 LVVFLVPLCLPVMMLNEGAEFCMVLVMAIFWVTEALPLVYTSMTPIVAFPPINGIMS 85  
DB 15 LIVFVPIILLPLILVPSKEAYCAVAILMALFWCTEALPLAVTALPFLILFPMGIVD 74  
QY 86 SDQTCRLYFKDPLVFMFGIMVALAVEYCNLHKRLARVIOVGCSPRLHFLGIMVTF 145  
DB 75 ASEVAVEYKDSNLLFFGGLLVAFAVEHNLKRLARVLIIVGRPAPLPLGFMVTF 134  
QY 146 LSMWISNAACTAMMCPPIQAVLEELQAGVCKINHEPQYIVGNGKNKNEDEPPYTK-- 203  
DB 135 LSMWISNATSAVMVPIAHAVLDQLHSSQASNVVE-----GNNPTFELQEPSQKEV 188  
QY 204 -----ITLCYVILGIAYASSLGGCGTIGPATNLTPKG 235  
DB 189 TKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCMSLCVCYSASIGGIATLTGTAPNLVLQ 248

QY 236 IYBARFKNSTEQMDPFTFMFYSVPMSLVYTLTFVFLQHFHMGCL-WRPKSKAEQVQRGR 294  
Db 249 QINSLFPQNGNVNFAFWFSFAFPTTVILLLLAWLWQLFLGFNFRKNFGIGERKQ--- 305  
QY 295 EGADVAKKVIDQRYKDLGPMSEIHEIOVMILFIFMVMYFTRKPGIFLGWADLL--NSKD- 351  
Db 306 EQQAAVCVITQTEHRLGPMFTFAKAIISILFVILVLLWFTREPGFFLGWGNLAFPPNAKGE 365  
QY 352 --IRNSMPTTFVVMCFMPLPANYAFRLYCTRRGGPV--PTGPTPSLITWKFIQTKVPWGL 407  
Db 366 SMVSDGTVAIFIGIIMFIIPSKPGLTQDPENPGKLKAPLG---LLDWTNVQKMPWNI 421  
QY 408 VFLGGGFALAEGSKQSGMAKLIGNALIGLKVLPNSVLLVILVAVLTAFFSNVAIAN 467  
Db 422 VLLGGGYALAKGSGSLSEWLNKLTPLQSPAPAPAIILSLLVATFTECTSNVATTT 481  
QY 468 IIPVLAEMSLATEIHPYLVLPCTLATSALFMLPVATPPNAIVFSGDLKVLDMARAGF 541  
Db 482 IFDILASMAQACILHPLVYMLPCTLATSALFMLPVATPPNAIVFSGDLKVLDMARAGF 541  
QY 528 GPTIITITLTFVFCQTVGLVYVYVNLNSFPPEWAIYAAA----ALGNKT 571  
Db 542 LLNIIGVLIATLAINSGIPLF-SLHSPFSPSAQSNNTTAQCLPSLANTT 588  
RESULT 13  
ADE63755  
ID ADE63755 standard; protein, 592 AA.  
XX  
AC ADE63755;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human Protein Q13183, SEQ ID NO 9699.  
XX  
KW Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury;CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
PN WO2003016475-A2.  
XX  
XX 27-FEB-2003.  
PD  
XX 14-AUG-2002; 2002WO-US025765.  
PF  
XX 14-AUG-2001; 2001US-0312147P.  
PR  
XX 01-NOV-2001; 2001US-0346382P.  
PR  
XX 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GEO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX  
DR WPI; 2003-268312/26.  
DR GENBANK; Q13183.  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017pp; English.  
XX  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence

CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 592 AA;  
Query Match 34.3%; Score 1026; DB 7; Length 592;  
Best Local Similarity 37.2%; Pred. No. 4.4e-103;  
Matches 219; Conservative 124; Mismatches 189; Indels 56; Gaps 10;  
QY 26 LVVFLVPLLCPLVMLINEGAEFRCMYLLLVMAIFWVTEALPLYVTSMIPVAPFIMGIMS 85  
Db 15 LVVFFVPIILLPLPILVPSKEAYCAIILMALFWCTEALPLAVTALFPLIFPMMGIVD 74  
QY 86 SDQTCRLYFKDTLVMPMGIMVALAVEYCNLHKRLALRYQIVGCSPRILHGLIMVTF 145  
Db 75 ASEVAVEYLKDSNLLPFGGLLVAVIAVEHWNHKLALRYLVGVREPAPLILGFLMTAF 134  
QY 146 LSMWISNAACTAMKCPILQAVLEELQAGVCKINHEPQYIVGKNKNEDEBPYTK-- 203  
Db 135 LSMWISNTATSAMVPIAHAVLDQJHSSQASSNVE-----GSNNPTTELQEPSQKEV 198  
QY 204 -----ITLCYVLGIAYASSLGCGGTIIGTATNLTFFKG 235  
Db 189 TKLDNGQALPVTSASSEGRAHLSQKHLHLTQCMSLCVCSASIGGIATLTGTAPNLVLOG 248  
QY 236 IYBARFKNSTEQMDPFTFMFYSVPMSLVYTLTFVFLQHFHMGCL-WRPKSKAEQVQRGR 294  
Db 249 QINSLFPQNGNVNFAFWFSFAFPTTVILLLLAWLWQLFLGFNFRKNFGIGERKQ--- 305  
QY 295 EGADVAKKVIDQRYKDLGPMSEIHEIOVMILFIFMVMYFTRKPGIFLGWADLL--NSKD- 351  
Db 306 EQQAAVCVITQTEHRLGPMFTFAKAIISILFVILVLLWFTREPGFFLGWGNLAFPPNAKGE 365  
QY 352 --IRNSMPTTFVVMCFMPLPANYAFRLYCTRRGGPV--PTGPTPSLITWKFIQTKVPWGL 407  
Db 366 SMVSDGTVAIFIGIIMFIIPSKPGLTQDPENPGKLKAPLG---LLDWTNVQKMPWNI 421  
QY 408 VFLGGGFALAEGSKQSGMAKLIGNALIGLKVLPNSVLLVILVAVLTAFFSNVAIAN 467  
Db 422 VLLGGGYALAKGSGSLSEWLNKLTPLQSPAPAPAIILSLLVATFTECTSNVATTT 481  
QY 468 IIPVLAEMSLATEIHPYLVLPCTLATSALFMLPVATPPNAIVFSGDLKVLDMARAGF 541  
Db 482 IFDILASMAQACILHPLVYMLPCTLATSALFMLPVATPPNAIVFSGDLKVLDMARAGF 541  
QY 528 GPTIITITLTFVFCQTVGLVYVYVNLNSFPPEWAIYAAA----ALGNKT 571  
Db 542 LLNIIGVLIATLAINSGIPLF-SLHSPFSPSAQSNNTTAQCLPSLANTT 588  
RESULT 14  
ADG16995  
ID ADG16995 standard; protein; 619 AA.  
XX  
AC ADG16995;  
XX





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:36:57; Search time 21.9924 Seconds  
(without alignments)  
1941.542 Million cell updates/sec

Title: US-10-017-479A-2  
Perfect score: 2994  
Sequence: 1 MBIETGEQPPVKCSNFPA.....NSFPEWAIYAAAALGNKTH 572

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database: Issued Patents AA:  
1: /cgn2\_6/prodata/1/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1742	58.2	368	4	US-09-270-767-42029
2	1026	34.3	599	4	US-09-949-016-9866
3	898	30.0	169	4	US-09-270-767-57286
4	896	29.9	626	4	US-09-556-916-20
5	895.5	29.9	627	4	US-09-556-916-8
6	891	29.8	626	4	US-09-556-916-14
7	891	29.8	626	4	US-09-556-916-22
8	890.5	29.7	627	4	US-09-556-916-2
9	890.5	29.7	627	4	US-09-556-916-10
10	889.5	29.7	627	4	US-09-949-016-6840
11	886	29.6	626	4	US-09-556-916-16
12	885.5	29.6	627	4	US-09-556-916-4
13	879.5	29.4	601	4	US-09-949-016-9977
14	760	25.4	580	4	US-09-556-916-24
15	759.5	25.4	581	4	US-09-556-916-12
16	755	25.2	580	4	US-09-556-916-18
17	754.5	25.2	581	4	US-09-556-916-6
18	711	23.7	561	4	US-09-949-016-8161
19	605.5	20.2	527	4	US-09-602-787A-516
20	577	19.3	230	4	US-09-270-767-43713
21	498.5	16.6	194	4	US-09-270-767-59102
22	463	15.5	180	4	US-09-270-767-42669
23	457	15.3	470	4	US-09-543-681A-5952
24	397	13.3	233	4	US-09-489-847-176
25	362	12.1	132	4	US-09-270-767-57987
26	357	11.9	335	4	US-09-602-787A-518
27	344	11.5	524	3	US-09-134-001C-5457

28	258.5	8.6	421	4	US-09-248-796A-20749	Sequence 20749, A
29	187.5	6.3	624	4	US-09-543-681A-4343	Sequence 4343, Ap
30	179.5	6.0	548	4	US-09-902-540-11870	Sequence 11870, A
31	176.5	5.9	596	4	US-09-902-540-13547	Sequence 13547, A
32	172.5	5.8	493	4	US-09-540-236-2120	Sequence 2120, Ap
33	166.5	5.6	302	4	US-09-902-540-10445	Sequence 10445, A
34	162	5.4	470	4	US-09-438-185A-209	Sequence 209, App
35	159	5.3	430	3	US-09-134-001C-2981	Sequence 2981, Ap
36	150.5	5.0	547	4	US-09-489-039A-13843	Sequence 13843, A
37	147	4.9	443	4	US-09-602-787A-532	Sequence 532, App
38	137	4.6	167	4	US-09-248-796A-20740	Sequence 20740, A
39	137	4.6	363	4	US-09-270-767-42613	Sequence 42613, A
40	136.5	4.6	478	3	US-09-134-001C-4637	Sequence 4637, Ap
41	136.5	4.6	494	4	US-09-543-681A-7033	Sequence 7033, Ap
42	135	4.5	694	4	US-09-252-991A-22637	Sequence 22637, A
43	134.5	4.5	436	4	US-09-583-110-4729	Sequence 4729, Ap
44	134.5	4.5	440	4	US-09-107-433-5201	Sequence 5201, Ap
45	134	4.5	437	4	US-09-543-681A-6984	Sequence 6984, Ap
46	131.5	4.4	614	4	US-09-489-039A-12605	Sequence 12605, A
47	131	4.4	319	4	US-09-270-767-57927	Sequence 57927, A
48	123	4.1	449	4	US-09-328-352-7512	Sequence 7512, Ap
49	122	4.1	465	4	US-09-543-681A-5337	Sequence 5337, Ap
50	120	4.0	569	2	US-08-750-723A-2	Sequence 2, Appli
51	120	4.0	569	3	US-09-191-275-2	Sequence 2, Appli
52	116.5	3.9	323	4	US-09-489-039A-7408	Sequence 7408, Ap
53	115.5	3.9	370	4	US-09-902-540-10305	Sequence 10305, A
54	114.5	3.8	502	4	US-09-252-991A-23327	Sequence 23327, A
55	112	3.7	477	3	US-09-134-001C-3487	Sequence 3487, Ap
56	111	3.7	457	4	US-09-489-039A-12762	Sequence 12762, A
57	111	3.7	484	4	US-09-107-532A-3830	Sequence 3830, Ap
58	110	3.7	429	4	US-08-311-731A-287	Sequence 287, App
59	107	3.6	476	4	US-09-543-681A-5260	Sequence 5260, Ap
60	106	3.5	494	2	US-09-031-392-5	Sequence 5, Appli
61	106	3.5	494	3	US-09-299-549-5	Sequence 5, Appli
62	106	3.5	446	4	US-09-610-417-5	Sequence 5, Appli
63	105	3.5	446	4	US-09-252-991A-17185	Sequence 17185, A
64	105	3.5	451	4	US-09-328-352-5922	Sequence 5922, Ap
65	104	3.5	367	3	US-09-134-001C-5557	Sequence 5557, Ap
66	104	3.5	398	4	US-09-489-039A-13159	Sequence 13159, A
67	103.5	3.5	381	4	US-09-710-279-3284	Sequence 3284, Ap
68	103.5	3.5	394	4	US-09-134-000C-3732	Sequence 3732, Ap
69	102	3.4	116	4	US-09-602-787A-520	Sequence 520, App
70	102	3.4	478	4	US-09-134-000C-6219	Sequence 6219, Ap
71	102	3.4	487	4	US-09-489-039A-8887	Sequence 8887, Ap
72	102	3.4	527	4	US-09-252-991A-18859	Sequence 18859, A
73	102	3.4	1014	4	US-09-253-991A-29868	Sequence 29868, A
74	101.5	3.4	503	4	US-09-583-110-4277	Sequence 4277, Ap
75	101.5	3.4	536	4	US-09-107-433-2945	Sequence 2945, Ap
76	101	3.4	321	4	US-09-107-532A-4118	Sequence 4118, Ap
77	100	3.3	345	4	US-09-902-540-9836	Sequence 9836, Ap
78	99	3.3	211	4	US-09-198-452A-1166	Sequence 1166, Ap
79	99	3.3	349	4	US-09-902-540-13275	Sequence 13275, A
80	98.5	3.3	441	4	US-09-489-039A-10612	Sequence 10612, A
81	98	3.3	557	4	US-09-540-236-2206	Sequence 2206, Ap
82	97.5	3.3	444	4	US-09-603-208A-240	Sequence 240, App
83	96.5	3.2	447	3	US-09-134-001C-4523	Sequence 4523, Ap
84	96.5	3.2	474	4	US-09-489-039A-14147	Sequence 14147, A
85	96.5	3.2	630	3	US-09-134-001C-4615	Sequence 4615, Ap
86	96	3.2	514	4	US-08-252-991A-30952	Sequence 30952, A
87	95.5	3.2	573	4	US-09-042-709A-19	Sequence 19, Appli
88	95	3.2	279	4	US-09-134-000C-5300	Sequence 5300, Ap
89	95	3.2	466	4	US-09-438-185A-231	Sequence 231, App
90	95	3.2	511	4	US-09-679-686B-12	Sequence 12, Appli
91	95	3.2	713	4	US-09-543-681A-6094	Sequence 6094, Ap
92	94.5	3.2	404	4	US-08-328-352-6361	Sequence 6361, Ap
93	94.5	3.2	574	2	US-08-948-569A-4	Sequence 4, Appli
94	94.5	3.2	574	2	US-09-188-469-4	Sequence 4, Appli
95	94.5	3.2	574	3	US-09-397-238A-4	Sequence 4, Appli
96	94.5	3.2	635	2	US-09-014-969-11	Sequence 11, Appli
97	94	3.1	462	2	US-08-898-976-2	Sequence 2, Appli
98	94	3.1	462	2	US-08-898-976-4	Sequence 4, Appli
99	94	3.1	471	4	US-09-543-681A-5580	Sequence 5580, Ap
100	94	3.1	502	4	US-09-902-540-13393	Sequence 13393, A



QY 187 VCGNKKNEDEPPYPTKIITLCYVLGIAYASSLGGCGTIIIGTATNLTFFKIYIYARFKNSTE 246  
DB 1 VCGNKKNEDEPPYPTKIITLCYVLGIAYASSLGGCGTIIIGTATNLTFFKIYIYARFKNSTE 60  
QY 247 QMDPPTFMFYSVPSMLVYLLTFVFLQWHPMGLWRPKSKEAOEVQREGADVAKKVIDQ 306  
DB 61 QMDPPTFMFYSVPSMLVYLLTFVFLQWHPMGLWRPKSKEAOEVQREGADVAKKVIDQ 120  
QY 307 RYKDLGPMSEIHEIQVMILFIWMVYFTRKPGIFLGWADLLNSKDIRNS 355  
DB 121 RYKDLGPMSEIHEIQVMILFIWMVYFTRKPGIFLGWADLLNSKDIRNS 169

RESULT 4  
US-09-556-916-20  
; Sequence 20, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 20  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-556-916-20

Query Match 29.9%; Score 896; DB 4; Length 626;  
Best Local Similarity 32.5%; Pred. No. 8.9e-88;  
Matches 205; Conservative 122; Mismatches 188; Indels 116; Gaps 15;

QY 24 KGLVFLVPLLCPLVMLNEGAEFCMVLVMAIFWVTEALPLVYVTSIMPIVAFPIGMI 83  
DB 12 KLLLVVCPVLLLPVLPVHPSSCAVLIIVAVYVSEAVPLGAAALVPAFLYPPFGV 71  
QY 84 MSSDOTCLYFKDTLVMPMGIMVALAVEYCNLHKRLALRVIOIGVCSPRRLHFLGIMVT 143  
DB 72 LRSNEVAAYEYFNTTLLVGVICVAAVEKWNLHKRIALRMVLMAGAKPGMLLLCFMCT 131  
QY 144 MFLSMWISNACTAMWCPIIOAVLEELQAGVCKINHEPQYQIVGNGKKNNEDEP- 198  
DB 132 TLLSMWLSNTSTTAMVPIVEAVLOEL-----VSAEDE-QLVAGNSNTEAEPIISLDV 183  
QY 199 ----- 198

DB 184 KNSQPSLELIFVNEESNADLTTLHMENLNGVPSITNPIKTANQHQKQHPQSKPOVL 243  
QY 199 -PYPTK-----ITLCYVLGIAYASSLGGCGTIIIGTATNLTFFKIYIYARFKN 243  
DB 244 TPSPRKQKLNRYRSHHDQMICKLSLSYSATIGGLTTIIGTSTL-----IFLEHFN 299  
QY 244 ---STEQMDPPTFMFYSVPSMLVYLLTFVFLQWHPMGL-WRPKSKEAOEVQREGADV 299  
DB 300 QYPAAEVVNFVGFWSFPIISLIMLVSWFWMHMLFLGNCNFKETCSLSKSKKTKRE--QL 357  
QY 300 AKKVIDORYKDLGPMSEIHEIQVMILFIWMVYFTRKPGIFLGWADLLNSKDIR-NSMPT 358  
DB 358 SEKRLOEYKELGDISYEMVTGFFILMTVLFTRFPGVPGWDFPEKKGYRTDATVS 417  
QY 359 IFVVMVCMFLPANYAFRLYCTRGGVPVPTGP-----TPSLITWKFIQTKVPMGLVFL 411  
DB 418 VFLGLFLFLIPAK-----KPCF---GKNDGENQEHSLGTESIITWKDFQKTMPEIVILV 470  
QY 412 GGGFALAGSKSQSGMAKLGIALGKVLPN-SVLLLVVILVAVFLTAFSSNVAIANII 470

DB 471 GGGYALASGSKSGSLSTWIGNQMLSSLPPWAVTLLACILVSI-VTFEVSNPATITIFL 529  
QY 471 PVLAEMSLAIEHPILYLLPAGLACSMAPHLFVSTPPNALVAGYANIRTKDMAIAGIGPT 530  
DB 530 PILCSLSETLHINPLYTLIPVTCISFAVMLPVGNPPNAIVPSYGHCIKDMVKAGLGVN 589  
QY 531 IITITLTFVFCOTWGLVVPYNLNSPPEWAOI 561  
DB 590 VIGLVVMVAINTMGVSLEF-HLDTYPAWARV 619

RESULT 5  
US-09-556-916-8  
; Sequence 8, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-556-916-8

Query Match 29.9%; Score 895.5; DB 4; Length 627;  
Best Local Similarity 32.4%; Pred. No. 1e-87;  
Matches 205; Conservative 122; Mismatches 188; Indels 117; Gaps 15;

QY 24 KGLVFLVPLLCPLVMLNEGAEFCMVLVMAIFWVTEALPLVYVTSIMPIVAFPIGMI 83  
DB 12 KLLLVVCPVLLLPVLPVHPSSCAVLIIVAVYVSEAVPLGAAALVPAFLYPPFGV 71  
QY 84 MSSDOTCLYFKDTLVMPMGIMVALAVEYCNLHKRLALRVIOIGVCSPRRLHFLGIMVT 143  
DB 72 LRSNEVAAYEYFNTTLLVGVICVAAVEKWNLHKRIALRMVLMAGAKPGMLLLCFMCT 131  
QY 144 MFLSMWISNACTAMWCPIIOAVLEELQAGVCKINHEPQYQIVGNGKKNNEDEP- 198  
DB 132 TLLSMWLSNTSTTAMVPIVEAVLOEL-----VSAEDE-QLVAGNSNTEAEPIISLDV 183  
QY 199 ----- 198

DB 184 KNSQPSLELIFVNEEDRSNADLTTLHMENLNGVPSITNPIKTANQHQKQHPQSKPOV 243  
QY 199 -PYPTK-----ITLCYVLGIAYASSLGGCGTIIIGTATNLTFFKIYIYARFKN 242  
DB 244 LTPSPRKQKLNRYRSHHDQMICKLSLSYSATIGGLTTIIGTSTL-----IFLEHFN 299  
QY 243 N---STEQMDPPTFMFYSVPSMLVYLLTFVFLQWHPMGL-WRPKSKEAOEVQREGAD 298  
DB 300 NOYPAAEVVNFVGFWSFPIISLIMLVSWFWMHMLFLGNCNFKETCSLSKSKKTKRE--Q 357  
QY 299 VAKKVIDORYKDLGPMSEIHEIQVMILFIWMVYFTRKPGIFLGWADLLNSKDIR-NSMP 357  
DB 358 LSEKRIQEEYKELGDISYEMVTGFFILMTVLFTRFPGVPGWDFPEKKGYRTDATV 417  
QY 358 TIFVVMVCMFLPANYAFRLYCTRGGVPVPTGP-----TPSLITWKFIQTKVPMGLVFL 410  
DB 418 SVFLGLFLFLIPAK-----KPCF---GKNDGENQEHSLGTESIITWKDFQKTMPEIVIL 470  
QY 411 LGGGFALAGSKSQSGMAKLGIALGKVLPN-SVLLLVVILVAVFLTAFSSNVAIANII 469  
DB 471 VGGYALASGSKSGSLSTWIGNQMLSSLPWAVTLLACILVSI-VTFEVSNPATITIF 529

QY 470 IPVLAEMSLAIEIHPYLLIPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIP 529  
Db 530 LPILCSLSETHINPLTYLIPVTMCISFAVMLPVGNNPNAIVFSYGHQCIKDMVKAGLV 589  
QY 530 TIITITLTFVFCOTWGLVVPNLNSPPEWAOI 561  
Db 590 NVIGLVIMVAINTWGVSLF-HLDTYPAWARV 620

RESULT 6  
US-09-556-916-14  
; Sequence 14, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; FILE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-556-916-14

Query Match 29.8%; Score 891; DB 4; Length 626;  
Best Local Similarity 32.3%; Pred. No. 3.1e-87;  
Matches 204; Conservative 122; Mismatches 189; Indels 116; Gaps 15;

QY 24 KGLVFLVPLLCPLVMLNEGAEFCRMVLLVMAIFWVTEALPLYVTSMIPVAFPIMG 83  
Db 12 KLLLVVCVPLLLPLVLPSPSEASCAVLIIVTVVWSEAVPLGAALVPAFLYPPFGV 71  
QY 84 MSSDOTCLRYFKDTLVMPMGIMVALAVEYCNLHKLRLALRVIQVGCSPRRHLHGLIMVT 143  
Db 72 LRSNEVAAYEFKNTLLLVGVICVAAAVEKNLHKLRLALRVLVMAKAPGMLLFCMCCT 131  
QY 144 MFLSMWISNACTAMWCPIQAVLEELQAGVCKINHEPOQIVGGNKKNEDEP----- 198  
Db 132 TLLSMWLSNTSTTAMVPIVEAVLQEL-----VSAEDE-QLVAGNSNTEEAEPISLDV 183  
QY 199 ----- 198  
Db 184 KNSQPSLELIFVNEESNADLTTLMHENLNGVPSITNPIKTANQHOGKKQHPSQEKPQVL 243  
QY 199 -PYPTK-----ITLCYVLGIAYASSLGGCGTIICTATNLTFKGIYEAFKN 243  
Db 244 TPSPRKQKLNRYSHHDDQICKLSLSISYSATIGGLTIIGTSTL-----IFLEHFN 299  
QY 244 ---STEQMDPTFMFYSPVSMVYLLTFVFLQWHEMGL-WRPKSKAEQVQREGADV 299  
Db 300 QYPAAEVNFGFWLFPSPISLIMLVSWFWHMLFLGCNFKETCSLSKKKTKRE--QL 357  
QY 300 AKKVIDORYKDLGPMSEIHEIQVMILFIFVMVYFTRKPGIFLGWADLLNSKDRI-NSMPT 358  
Db 358 SEKRIOEYKLGDISYPEWVTGFFILMTVLTFTREPFGVPGWDSFPEKKGYRTDATVS 417  
QY 359 IFVVMVCMFLPANYAFRLYCTRGGPVPTGP-----TPSLITWKFIQTKVPMGLVFL 411  
Db 418 VFLGFLFLIPAK-----KPCF---GKNDGENQEHSLGTETITWKDFQKTMPEIVILV 470  
QY 412 GGGFALAECSKSGMAKIGNALIGLVLPN-SVLLLVILVAVPLTAFSSNVAIANII 470  
Db 471 GGGYALASGSKSGSLTWIGNQMLSSLPWAVTLLACILVSI-VTFEVSNPATITIFL 529  
QY 471 PVLAEMSLAIEIHPYLLIPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIPT 530  
Db 530 LPILCSLSETHINPLTYLIPVTMCISFAVMLPVGNNPNAIVFSYGHQCIKDMVKAGLV 589

Db 530 PILCSLSETHINPLTYLIPVTMCISFAVMLPVGNNPNAIVFSYGHQCIKDMVKAGLVN 589  
QY 531 ITITITLTFVFCOTWGLVVPNLNSPPEWAOI 561  
Db 590 NVIGLVIMVAINTWGVSLF-HLDTYPAWARV 619

RESULT 7  
US-09-556-916-22  
; Sequence 22, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; FILE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 22  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-556-916-22

Query Match 29.8%; Score 891; DB 4; Length 626;  
Best Local Similarity 32.5%; Pred. No. 3.1e-87;  
Matches 205; Conservative 121; Mismatches 189; Indels 116; Gaps 15;

QY 24 KGLVFLVPLLCPLVMLNEGAEFCRMVLLVMAIFWVTEALPLYVTSMIPVAFPIMG 83  
Db 12 KLLLVVCVPLLLPLVLPSPSEASCAVLIIVTVVWSEAVPLGAALVPAFLYPPFGV 71  
QY 84 MSSDOTCLRYFKDTLVMPMGIMVALAVEYCNLHKLRLALRVIQVGCSPRRHLHGLIMVT 143  
Db 72 LRSNEVAAYEFKNTLLLVGVICVAAAVEKNLHKLRLALRVLVMAKAPGMLLFCMCCT 131  
QY 144 MFLSMWISNACTAMWCPIQAVLEELQAGVCKINHEPOQIVGGNKKNEDEP----- 198  
Db 132 TLLSMWLSNTSTTAMVPIVEAVLQEL-----VSAEDE-QLVAGNSNTEEAEPISLDV 183  
QY 199 ----- 198  
Db 184 KNSQPSLELIFVNEESNADLTTLMHENLNGVPSITNPIKTANQHOGKKQHPSQEKPQVL 243  
QY 199 -PYPTK-----ITLCYVLGIAYASSLGGCGTIICTATNLTFKGIYEAFKN 243  
Db 244 TPSPRKQKLNRYSHHDDQICKLSLSISYSATIGGLTIIGTSTL-----IFLEHFN 299  
QY 244 ---STEQMDPTFMFYSPVSMVYLLTFVFLQWHEMGL-WRPKSKAEQVQREGADV 299  
Db 300 QYPAAEVNFGFWLFPSPISLIMLVSWFWHMLFLGCNFKETCSLSKKKTKRE--QL 357  
QY 300 AKKVIDORYKDLGPMSEIHEIQVMILFIFVMVYFTRKPGIFLGWADLLNSKDRI-NSMPT 358  
Db 358 SEKRIOEYKLGDISYPEWVTGFFILMTVLTFTREPFGVPGWDSFPEKKGYRTDATVS 417  
QY 359 IFVVMVCMFLPANYAFRLYCTRGGPVPTGP-----TPSLITWKFIQTKVPMGLVFL 411  
Db 418 VFLGFLFLIPAK-----KPCF---GKNDGENQEHSLGTETITWKDFQKTMPEIVILV 470  
QY 412 GGGFALAECSKSGMAKIGNALIGLVLPN-SVLLLVILVAVPLTAFSSNVAIANII 470  
Db 471 GGGYALASGSKSGSLTWIGNQMLSSLPWAVTLLACILVSI-VTFEVSNPATITIFL 529  
QY 471 PVLAEMSLAIEIHPYLLIPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIPT 530  
Db 530 LPILCSLSETHINPLTYLIPVTMCISFAVMLPVGNNPNAIVFSYGHQCIKDMVKAGLVN 589

QY 531 IITITITLVFCQWGLVVPVNLNSFPFPAQI 561  
Db 590 VIGLVIVMVAINTWGVSLF-HLDTYPAWARV 619

## RESULT 8

US-09-556-916-2

; Sequence 2, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-556-916-2

Query Match 29.7%; Score 890.5; DB 4; Length 627;  
Best Local Similarity 32.3%; Pred. No. 3.5e-87;  
Matches 204; Conservative 122; Mismatches 189; Indels 117; Gaps 15;

QY 24 KGLVVLVPLLCPLVMLNEGAFRCMYLLLVMAIFVWTEALPLVVTSMIPVAFPIGMI 83  
Db 12 KLLLVVCPVLLLLPLVLPVSPSEASCAVYLVITAVVWSEAVPLGAAALVPAFLVPPFGV 71  
QY 84 MSSDOTCRLYFKDTLVMPFGGIMVALAVEYCNLHKRLALRVIOVGCSPRLHFGLIWVT 143  
Db 72 LRSNEVAAYEYFNKNTLLLVGVICVAAAVEKNLHKRIALRMVLMAKPGMILLCFMCT 131  
QY 144 MFLSNWISNAACTAMWCPITQAVLEEQAOQGVCKINHEPOYIVGNGKNKNEDEP- 198  
Db 132 TLLSNWLSNTSTAMVPIVEAVLQEL-----VSAEDE-QLVAGNSNTEAEPISLDV 183  
QY 199 ----- 198  
Db 184 KNSQPSLELIFVNEEDRSNADLTTLNHNENLNGVPSITNPIKTANQHQKQHPQV 243  
QY 199 --PYTK-----ITLCYVIGIAYASSLGGCGTIIIGTATNLTFGIYEAPK 242  
Db 244 LTPSPRKQKLNRYRSHDDQMICLSLSISYSATIGGLTIIIGTSTSL-----IFLEHN 299  
QY 243 N---STEQMDRPTFMFYSVPSMLVYLLTFVFLQWHEMGL-WRPKSKAEQVQREGAD 298  
Db 300 NOYPAAEVNFGTWFSPISLIMLVSWFWMMHLFLGCFKTCSLSKKKTKRE--Q 357  
QY 299 VAKKVIDORYKDLGPMSEIHEIQVMILFIMVVMYFTRKPGIFLGWADLLNSKDIR-NSMP 357  
Db 358 LSEKIQEYKGLGDISPEWMTGFFILMTVLMFTREPFGVPGWDSFFEKKGYRTDATV 417  
QY 358 TIFVVMCFMPLPANYAFIRYCTRGGVPVTPG-----TPSLITWKFQITKVPGLVFL 410  
Db 418 SVFLGFLFLIPAK----KPCF---GKNDGENQHSLSGTETPIITWKDFQKTMPEVIL 470  
QY 411 LGGCFALAGSKQSGMAKLGNALIGLVLPN-SVLLLVILVAVFLTAFSSNAIANII 469  
Db 471 VGGYALAGSKSGSLSTWIGNQMSLSSLPWAVTLACILVSI-VTEFVSNPATITIF 529  
QY 470 IPVLAEMSLAIEIHPYLLILPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIGIP 529  
Db 530 LPILCSLSETHINPLTYLIPVTWCISFAVMLPVGNPNNAIVFSYGHQCIQKMWKAGLV 589  
QY 530 TIITITITLVFCQWGLVVPVNLNSFPFPAQI 561  
Db 590 VIGLVIVMVAINTWGVSLF-HLDTYPAWARV 619

Db 590 NVIGLVIVMVAINTWGVSLF-HLDTYPAWARV 620

## RESULT 9

US-09-556-916-10

; Sequence 10, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 10  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-556-916-10

Query Match 29.7%; Score 890.5; DB 4; Length 627;  
Best Local Similarity 32.4%; Pred. No. 3.5e-87;  
Matches 205; Conservative 121; Mismatches 189; Indels 117; Gaps 15;

QY 24 KGLVVLVPLLCPLVMLNEGAFRCMYLLLVMAIFVWTEALPLVVTSMIPVAFPIGMI 83  
Db 12 KLLLVVCPVLLLLPLVLPVSPSEASCAVYLVITAVVWSEAVPLGAAALVPAFLVPPFGV 71  
QY 84 MSSDOTCRLYFKDTLVMPFGGIMVALAVEYCNLHKRLALRVIOVGCSPRLHFGLIWVT 143  
Db 72 LRSNEVAAYEYFNKNTLLLVGVICVAAAVEKNLHKRIALRMVLMAKPGMILLCFMCT 131  
QY 144 MFLSNWISNAACTAMWCPITQAVLEEQAOQGVCKINHEPOYIVGNGKNKNEDEP- 198  
Db 132 TLLSNWLSNTSTAMVPIVEAVLQEL-----VSAEDE-QLVAGNSNTEAEPISLDV 183  
QY 199 ----- 198  
Db 184 KNSQPSLELIFVNEEDRSNADLTTLNHNENLNGVPSITNPIKTANQHQKQHPQV 243  
QY 199 --PYTK-----ITLCYVIGIAYASSLGGCGTIIIGTATNLTFGIYEAPK 242  
Db 244 LTPSPRKQKLNRYRSHDDQMICLSLSISYSATIGGLTIIIGTSTSL-----IFLEHN 299  
QY 243 N---STEQMDRPTFMFYSVPSMLVYLLTFVFLQWHEMGL-WRPKSKAEQVQREGAD 298  
Db 300 NOYPAAEVNFGTWFSPISLIMLVSWFWMMHLFLGCFKTCSLSKKKTKRE--Q 357  
QY 299 VAKKVIDORYKDLGPMSEIHEIQVMILFIMVVMYFTRKPGIFLGWADLLNSKDIR-NSMP 357  
Db 358 LSEKIQEYKGLGDISPEWMTGFFILMTVLMFTREPFGVPGWDSFFEKKGYRTDATV 417  
QY 358 TIFVVMCFMPLPANYAFIRYCTRGGVPVTPG-----TPSLITWKFQITKVPGLVFL 410  
Db 418 SVFLGFLFLIPAK----KPCF---GKNDGENQHSLSGTETPIITWKDFQKTMPEVIL 470  
QY 411 LGGCFALAGSKQSGMAKLGNALIGLVLPN-SVLLLVILVAVFLTAFSSNAIANII 469  
Db 471 VGGYALAGSKSGSLSTWIGNQMSLSSLPWAVTLACILVSI-VTEFVSNPATITIF 529  
QY 470 IPVLAEMSLAIEIHPYLLILPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIGIP 529  
Db 530 LPILCSLSETHINPLTYLIPVTWCISFAVMLPVGNPNNAIVFSYGHQCIQKMWKAGLV 589  
QY 530 TIITITITLVFCQWGLVVPVNLNSFPFPAQI 561  
Db 590 NVIGLVIVMVAINTWGVSLF-HLDTYPAWARV 620

```
RESULT 10
US-09-949-016-6840
; Sequence 6840, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6840
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6840

Query Match      29.7%; Score 889.5; DB 4; Length 627;
Best Local Similarity 32.3%; Pred. No. 4.5e-87;
Matches 204; Conservative 122; Mismatches 189; Indels 117; Gaps 15;

QY 24 KGLVFLVPLLCPLVMLNEGAFRCMYLLLVMAIFWVTEALPLVYVTSMIPVAFPIGMI 83
DB 12 KLLLVVCVPLLLPLVLPVHPSSSEASCAYLVITAVYVWSEAVPLGAALVPAFLYPPFGV 71
QY 84 MSSDQTCRLYFKDTLVPMFGIMVALAVEYCNLHKRLALRVIOIVGCSRRLLHFLGIMVT 143
DB 72 LRSNEVAAYEFKNTLLLVGVICVAAAVEKNWLHKRIALRMVLMAKAPGMLLLCFMCC 131
QY 144 MFLSMWISNACTAMMCPITQAVLEELQAGVCKINHEPOYQIVGKNKNEDEP----- 198
DB 132 TLLSMWLSNTSTTAVMPIVEAVLOEL-----VSAEDE-QLVAGNSNTEAEPISLDV 183
QY 199 -----VSAEDE-QLVAGNSNTEAEPISLDV 183
DB 184 KNSQPSLELIFVNEEDSNADLTTLHMNENLNGVPSITNPIKTANQHQKQHPQKQV 243
QY 199 --PYPTK-----ITLCYVLGIAYASSLGGCGTIIIGTATNLTPKGIYEARFK 242
DB 244 LTPSPRKQKLNRYRSHHDQMIKCLSLISYSATIGLTLTIIGTSTL-----IFLHFN 299
QY 243 N---STEQMDPPTFMFYSVPVSMVYVTLTFVFLQWHEMGL-WRPKSKAEQVQREGAD 298
DB 300 NQYPAAEVYVNGTWFPLSFPSISLIMLVVSWFWMHVFLGCFKETSLSKSKKTKRE--Q 357
QY 299 VAKKVIQRYKDLGPMSTHEIQVMILFTFMVVMYFTRKPGIFLGMADLLNSKDIR-NSMP 357
DB 358 LSEKRIQEEYEKLDGISYPMVMTGFFILMTLVMTFTREPFGVPCWDSFFEKKGYRTDATV 417
QY 358 TIFVVMCFMPLPANYAFILRYCTRGGVPVTPG-----TPSLITWKPIQTKVPMGLVFL 410
DB 418 SVFLGFLFLFLIPAK-----KPCF---GKKNNGENQEHSLGTEPIITWKDFQKTMPEVIL 470
QY 411 LGGFFALAEKSGKSMAGLIGNALIGLVLPN-SVLLLVLLVAVLFTATSSNVAIANII 469
DB 471 VGGFYALASGSKSGLSTWIGNQMLSSLPWPVAVTLTLLACILVSI-VTEFVSNPATITIF 529
QY 470 IPVLAEMSLAIEIHPVLLIPAGLACSMAPHLVSTTPNALVAGYANIRTKDMAIAGIGP 529
DB 530 LPILCSLSETHINPLTYLIPVTCISFVAVMLPVGNPNNAIVFSYGHCIQKDMVKAGLV 589
QY 530 TIITITLTFVPCQWGLVYVYVNLNSFPPEAQI 561
DB 590 NVIGLVIMVAINTWGVSLF-HLDTPAWARV 620
```

```
RESULT 11
US-09-556-916-16
; Sequence 16, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271 Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-16
```

```
Query Match      29.6%; Score 886; DB 4; Length 626;
Best Local Similarity 32.3%; Pred. No. 1.1e-86;
Matches 204; Conservative 121; Mismatches 190; Indels 116; Gaps 15;

QY 24 KGLVFLVPLLCPLVMLNEGAFRCMYLLLVMAIFWVTEALPLVYVTSMIPVAFPIGMI 83
DB 12 KLLLVVCVPLLLPLVLPVHPSSSEASCAYLVITAVYVWSEAVPLGAALVPAFLYPPFGV 71
QY 84 MSSDQTCRLYFKDTLVPMFGIMVALAVEYCNLHKRLALRVIOIVGCSRRLLHFLGIMVT 143
DB 72 LRSNEVAAYEFKNTLLLVGVICVAAAVEKNWLHKRIALRMVLMAKAPGMLLLCFMCC 131
QY 144 MFLSMWISNACTAMMCPITQAVLEELQAGVCKINHEPOYQIVGKNKNEDEP----- 198
DB 132 TLLSMWLSNTSTTAVMPIVEAVLOEL-----VSAEDE-QLVAGNSNTEAEPISLDV 183
QY 199 -----VSAEDE-QLVAGNSNTEAEPISLDV 198
DB 184 KNSQPSLELIFVNEEDSNADLTTLHMNENLNGVPSITNPIKTANQHQKQHPQKQV 243
QY 199 --PYPTK-----ITLCYVLGIAYASSLGGCGTIIIGTATNLTPKGIYEARFK 243
DB 244 TSPSPRKQKLNRYRSHHDQMIKCLSLISYSATIGLTLTIIGTSTL-----IFLHFN 299
QY 244 ---STEQMDPPTFMFYSVPVSMVYVTLTFVFLQWHEMGL-WRPKSKAEQVQREGADV 299
DB 300 QYPAAEVYVNGTWFPLSFPSISLIMLVVSWFWMHVFLGCFKETSLSKSKKTKRE--QL 357
QY 300 AKKVIQRYKDLGPMSTHEIQVMILFTFMVVMYFTRKPGIFLGMADLLNSKDIR-NSMPT 358
DB 358 SEKRIQEEYEKLDGISYPMVMTGFFILMTLVMTFTREPFGVPCWDSFFEKKGYRTDATV 417
QY 359 IFVVMCFMPLPANYAFILRYCTRGGVPVTPG-----TPSLITWKPIQTKVPMGLVFL 411
DB 418 VFLGFLFLFLIPAK-----KPCF---GKKNNGENQEHSLGTEPIITWKDFQKTMPEVILV 470
QY 412 GGGFALAEKSGKSMAGLIGNALIGLVLPN-SVLLLVLLVAVLFTATSSNVAIANII 470
DB 471 GGGFYALASGSKSGLSTWIGNQMLSSLPWPVAVTLTLLACILVSI-VTEFVSNPATITIFL 529
QY 471 PVLAEMSLAIEIHPVLLIPAGLACSMAPHLVSTTPNALVAGYANIRTKDMAIAGIGPT 530
DB 530 PILCSLSETHINPLTYLIPVTCISFVAVMLPVGNPNNAIVFSYGHCIQKDMVKAGLVN 589
QY 531 IITITLTFVPCQWGLVYVYVNLNSFPPEAQI 561
DB 590 VIGLVIMVAINTWGVSLF-HLDTPAWARV 619
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:59:01 ; Search time 74.2245 Seconds  
(without alignments)  
2973.073 Million cell updates/sec

Title: US-10-017-479A-2  
Perfect score: 2994  
Sequence: 1 MEIEIGEPQPPVKCSNFFA.....NSFPEWAIYAAALGNKTH 572

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1722976 seqs, 385795295 residues

Total number of hits satisfying chosen parameters: 1722976

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Published Applications\_AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pdb.pcp.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pdb.pcp.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pdb.pcp.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pdb.pcp.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pdb.pcp.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pdb.pcp.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pdb.pcp.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pdb.pcp.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pdb.pcp.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pdb.pcp.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pdb.pcp.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pdb.pcp.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pdb.pcp.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pdb.pcp.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pdb.pcp.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pdb.pcp.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pdb.pcp.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pdb.pcp.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pdb.pcp.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pdb.pcp.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pdb.pcp.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pdb.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2994	100.0	572	14	US-10-167-994-2
2	2994	100.0	572	14	US-10-017-479-2
3	2994	100.0	572	17	US-10-718-359-2
4	1323	44.2	562	14	US-10-167-994-10
5	1042.5	34.8	587	13	US-10-173-519-5
6	1042.5	34.8	587	14	US-10-167-994-12
7	1042.5	34.8	587	17	US-10-718-359-13
8	1039	34.7	581	17	US-10-718-359-12
9	1026	34.3	592	14	US-10-167-994-3
10	1026	34.3	592	14	US-10-167-994-11
11	1010.5	33.8	619	9	US-09-729-094-4
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 10, Appl
					Sequence 5, Appli
					Sequence 12, Appl
					Sequence 13, Appl
					Sequence 12, Appl
					Sequence 3, Appli
					Sequence 11, Appl
					Sequence 4, Appli

Sequence 4, Appli  
Sequence 2, Appli  
Sequence 72, Appl  
Sequence 1, Appli  
Sequence 6, Appli  
Sequence 44, Appl  
Sequence 70, Appl  
Sequence 149, Appl  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 46, Appl  
Sequence 48, Appl  
Sequence 68, Appl  
Sequence 4, Appli  
Sequence 13, Appl  
Sequence 14, Appl  
Sequence 6, Appli  
Sequence 4, Appli  
Sequence 25, Appl  
Sequence 23, Appl  
Sequence 10, Appl  
Sequence 20, Appl  
Sequence 8, Appli  
Sequence 14, Appl  
Sequence 22, Appl  
Sequence 10, Appl  
Sequence 7, Appli  
Sequence 16, Appl  
Sequence 4, Appli  
Sequence 118, Appl  
Sequence 5, Appli  
Sequence 9, Appli  
Sequence 50, Appl  
Sequence 74, Appl  
Sequence 362, Ap  
Sequence 52, Appl  
Sequence 76, Appl  
Sequence 8, Appli  
Sequence 24, Appl  
Sequence 12, Appl  
Sequence 18, Appl  
Sequence 6, Appli  
Sequence 436, App  
Sequence 3754, Ap  
Sequence 516, App  
Sequence 127443,  
Sequence 160, App  
Sequence 1947, Ap  
Sequence 250656,  
Sequence 127444,  
Sequence 346, App  
Sequence 1949, Ap  
Sequence 2022, Ap  
Sequence 21334, A  
Sequence 2634, A  
Sequence 176, App  
Sequence 7, Appli  
Sequence 354, App  
Sequence 354, App  
Sequence 354, App  
Sequence 354, App  
Sequence 354, App  
Sequence 254988,  
Sequence 518, App  
Sequence 4, Appli  
Sequence 584, App  
Sequence 7011, Ap  
Sequence 8, Appli  
Sequence 9336, Ap  
Sequence 500, App  
Sequence 5397, Ap  
Sequence 1187, Ap

US-10-435-631-4  
US-10-173-519-2  
US-10-403-161-72  
US-10-490-080-1  
US-10-718-359-6  
US-10-092-900A-44  
US-10-403-161-70  
US-10-128-558-149  
US-09-729-094-2  
US-10-435-631-2  
US-10-092-900A-46  
US-10-092-900A-48  
US-10-403-161-68  
US-10-718-359-4  
US-10-167-994-13  
US-10-718-359-14  
US-10-167-994-6  
US-10-167-994-4  
US-10-343-903-25  
US-10-490-080-23  
US-10-718-359-10  
US-10-368-687-20  
US-10-368-687-8  
US-10-368-687-14  
US-10-368-687-22  
US-10-368-687-10  
US-10-167-994-7  
US-10-368-687-16  
US-10-368-687-4  
US-10-034-934-118  
US-10-167-994-5  
US-10-380-727-9  
US-10-092-900A-50  
US-10-403-161-74  
US-10-108-260A-382  
US-10-092-900A-52  
US-10-403-161-76  
US-10-718-359-8  
US-10-368-687-24  
US-10-368-687-12  
US-10-368-687-6  
US-09-882-227-496  
US-09-738-626-3754  
US-10-627-476-516  
US-10-437-963-127443  
US-10-828-992-160  
US-09-833-245-1947  
US-10-425-115-250656  
US-10-437-963-127444  
US-10-128-558-346  
US-09-833-245-1949  
US-10-264-237-2022  
US-10-369-493-21334  
US-10-276-774-2634  
US-10-351-334-176  
US-10-173-519-7  
US-09-974-879-354  
US-09-305-736-354  
US-09-818-683-354  
US-09-818-683-354  
US-10-621-401-354  
US-10-424-599-254988  
US-10-627-476-518  
US-10-173-519-4  
US-10-470-048B-584  
US-10-741-849-7011  
US-10-173-519-8  
US-10-369-493-9936  
US-10-238-075-500  
US-10-335-977-5397  
US-10-296-115-1187

85 261 8.7 162 11 US-09-833-245-1948 Sequence 1948, Ap  
86 260.5 8.7 163 15 US-10-424-599-184881 Sequence 184881,  
87 258 8.6 236 16 US-10-767-701-43427 Sequence 43427, A  
88 257.5 8.6 162 11 US-09-833-245-1945 Sequence 1945, Ap  
89 241 8.0 259 16 US-10-739-930-9458 Sequence 9458, Ap  
90 234 7.8 226 15 US-10-296-115-980 Sequence 980, App  
91 232.5 7.8 300 16 US-10-767-701-46317 Sequence 46317, A  
92 228.5 7.6 174 16 US-10-425-115-335649 Sequence 335649,  
93 228 7.6 78 9 US-09-864-761-34252 Sequence 34252, A  
94 222 7.6 78 14 US-10-029-386-28038 Sequence 28038, A  
95 222 7.4 457 15 US-10-369-493-4266 Sequence 4266, Ap  
96 217 7.2 135 15 US-10-276-774-1592 Sequence 1592, Ap  
97 211 7.0 164 15 US-10-424-599-205152 Sequence 205152,  
98 185 6.2 482 15 US-10-335-977-5881 Sequence 5881, Ap  
99 184 6.1 487 15 US-10-369-493-23559 Sequence 23559, A  
100 179 6.0 494 15 US-10-335-977-5882 Sequence 5882, Ap

## ALIGNMENTS

## RESULT 1

US-10-167-994-2  
; Sequence 2, Application US/10167994  
; Publication No. US20030082647A1  
; GENERAL INFORMATION:  
; APPLICANT: Reenan, Robert A.  
; APPLICANT: Rogina, Blanka  
; APPLICANT: Helfand, Stephen L.  
; TITLE OF INVENTION: TRANSPORTER PROTEIN  
; FILE REFERENCE: 13407-013001  
; CURRENT APPLICATION NUMBER: US/10/167,994  
; CURRENT FILING DATE: 2002-06-12  
; PRIOR APPLICATION NUMBER: US 60/255,013  
; PRIOR FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: US 10/017,479  
; PRIOR FILING DATE: 2001-12-12  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 572  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-167-994-2

Query Match 100.0%; Score 2994; DB 14; Length 572;  
Best Local Similarity 100.0%; Pred. No. 2e-284;  
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 METEIGEQPPVKCSNFFANHWKGLVFLVPLLCPLVMLLNEGAERFCMYYLLLVMAIFW 60  
DB 1 METEIGEQPPVKCSNFFANHWKGLVFLVPLLCPLVMLLNEGAERFCMYYLLLVMAIFW 60  
QY 61 VTEALPLYVTSMIPVAFPIGIMSSDQTCRLYFKDTLVNFMGGMVVALAVEYCNLHKRL 120  
DB 61 VTEALPLYVTSMIPVAFPIGIMSSDQTCRLYFKDTLVNFMGGMVVALAVEYCNLHKRL 120  
QY 121 ALRVIQVGCSPRRHLHGLIMVTMFLSMWISNAACTAMMCPPIQAVLEELQAOGVCKINH 180  
DB 121 ALRVIQVGCSPRRHLHGLIMVTMFLSMWISNAACTAMMCPPIQAVLEELQAOGVCKINH 180  
QY 181 EPQYQIVGGKKNNEDEPPYPTKITLCYILGIAYASSLGGCGTIIIGTATNLTTPKGIYEAR 240  
DB 181 EPQYQIVGGKKNNEDEPPYPTKITLCYILGIAYASSLGGCGTIIIGTATNLTTPKGIYEAR 240  
QY 241 FKNSTEQMDPFTFMFYSPVSMYVTLTTFVFLQWHFMGLWRPKSKEAQEVQREGADVA 300  
DB 241 FKNSTEQMDPFTFMFYSPVSMYVTLTTFVFLQWHFMGLWRPKSKEAQEVQREGADVA 300  
QY 301 KKVIDORYKDLGPMSEIHEIQVMILFIMVMYFTRKPGIFLGWADLLNSKDINSMTPIF 360  
DB 301 KKVIDORYKDLGPMSEIHEIQVMILFIMVMYFTRKPGIFLGWADLLNSKDINSMTPIF 360  
QY 361 VVVMCFMFLPANTAFRLYCTRRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420  
DB 361 VVVMCFMFLPANTAFRLYCTRRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420  
QY 421 SKQSGMAKLIGNALIGLKVLPSNVLVAVFLTAFSSNVAIANIIPVLAEMSLAI 480

## RESULT 2

US-10-017-479-2  
; Sequence 2, Application US/10017479  
; Publication No. US20030104399A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Connecticut  
; APPLICANT: Helfand, Stephan L  
; APPLICANT: Reenan, Robert A  
; APPLICANT: Rogina, Blanka  
; TITLE OF INVENTION: Polynucleotides Encoding Cellular Transporters and Methods of Use  
; FILE REFERENCE: UCT-0020  
; CURRENT APPLICATION NUMBER: US/10/017,479  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: 60/255,013  
; PRIOR FILING DATE: 2000-12-12  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 2  
; LENGTH: 572  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-017-479-2

Query Match 100.0%; Score 2994; DB 14; Length 572;  
Best Local Similarity 100.0%; Pred. No. 2e-284;  
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 METEIGEQPPVKCSNFFANHWKGLVFLVPLLCPLVMLLNEGAERFCMYYLLLVMAIFW 60  
DB 1 METEIGEQPPVKCSNFFANHWKGLVFLVPLLCPLVMLLNEGAERFCMYYLLLVMAIFW 60  
QY 61 VTEALPLYVTSMIPVAFPIGIMSSDQTCRLYFKDTLVNFMGGMVVALAVEYCNLHKRL 120  
DB 61 VTEALPLYVTSMIPVAFPIGIMSSDQTCRLYFKDTLVNFMGGMVVALAVEYCNLHKRL 120  
QY 121 ALRVIQVGCSPRRHLHGLIMVTMFLSMWISNAACTAMMCPPIQAVLEELQAOGVCKINH 180  
DB 121 ALRVIQVGCSPRRHLHGLIMVTMFLSMWISNAACTAMMCPPIQAVLEELQAOGVCKINH 180  
QY 181 EPQYQIVGGKKNNEDEPPYPTKITLCYILGIAYASSLGGCGTIIIGTATNLTTPKGIYEAR 240  
DB 181 EPQYQIVGGKKNNEDEPPYPTKITLCYILGIAYASSLGGCGTIIIGTATNLTTPKGIYEAR 240  
QY 241 FKNSTEQMDPFTFMFYSPVSMYVTLTTFVFLQWHFMGLWRPKSKEAQEVQREGADVA 300  
DB 241 FKNSTEQMDPFTFMFYSPVSMYVTLTTFVFLQWHFMGLWRPKSKEAQEVQREGADVA 300  
QY 301 KKVIDORYKDLGPMSEIHEIQVMILFIMVMYFTRKPGIFLGWADLLNSKDINSMTPIF 360  
DB 301 KKVIDORYKDLGPMSEIHEIQVMILFIMVMYFTRKPGIFLGWADLLNSKDINSMTPIF 360  
QY 361 VVVMCFMFLPANTAFRLYCTRRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420  
DB 361 VVVMCFMFLPANTAFRLYCTRRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420  
QY 421 SKQSGMAKLIGNALIGLKVLPSNVLVAVFLTAFSSNVAIANIIPVLAEMSLAI 480



RESULT 5

US-10-173-519-5

; Sequence 5, Application US/10173519

; Publication No. US20020193582A1

; GENERAL INFORMATION:

; APPLICANT: Curtiss, Rory A.J.

; APPLICANT: Millennium Pharmaceuticals, Inc.

; TITLE OF INVENTION: 69624, A Human Transporter Family Member

; TITLE OF INVENTION: and Uses Therefor

; CURRENT APPLICATION NUMBER: US/10173519

; CURRENT FILING DATE: 2002-06-17

; PRIOR APPLICATION NUMBER: 60/298,970

; PRIOR FILING DATE: 2001-06-18

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 587

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-10-173-519-5

Query Match 34.8%; Score 1042.5; DB 13; Length 587;

Best Local Similarity 38.1%; Pred. No. 5.9e-93;

Matches 219; Conservative 129; Mismatches 190; Indels 37; Gaps 8;

QY 26 LVVFLVLLCLPVMNLNEGAFRCMYLLVMAIFWVTEALPLYVTSMIPVAPFIMGIMS 85  
 DB 15 LIVLCLPIFLPLPLIVQTKYACAYSIILMALLWCTEALPLAVTALFVILFPLMGIND 74  
 QY 86 SDQTCRLYFKDTLVFMWGGIMVALAVEYCNLHKLRLALRVQIVGCSRRLHFGILVMTWF 145  
 DB 75 ASEVCIEYFKDTNLFVGGIMVALAVEYCNLHKLRLALRVQIVGCSRRLHFGILVMTWF 134  
 QY 146 LSMWISNACTAMMCPITQAVLELQA--QGVCKINHEPOYQIV-----GGNK 191  
 DB 135 LSMWISNTATTAMVPIGHAVLEQLQSGKKDVEGNNNPTFELQEECPQKEVTKLDNGQP 194  
 QY 192 KNEDEP-----PYPTKITLCYVLGIAYASSIGCGGTIIGTATNLTFRGIYEAREFK 242  
 DB 195 VSAPSEPRTKTOEHRHFSQGLSLC-----ICYSASIGGIATLTGTPNLVLOGVNSLFP 250  
 QY 243 NSTEQMDPFTFMYSVPSMLVYLLTFVFLQWHFGLWRPKSEAEQVORGEGADVAKK 302  
 DB 251 QNGNVNFAWFGFAFTMIILLLLAWLQVLFLGVNFRKNFGFGEGERKQA--AFQ 308  
 QY 303 VIDORYKDLGPMSTHEIQVMILFIFWVYFTRKPGIFLGWADLLNSKDIRNSMPT---- 358  
 DB 309 VIKTQYRLGLGPMSPAEKTVTVLVLLVLFVTFREPFGFPFGWGTTFVANEKGQSMASDGTV 368  
 QY 359 -IFVVMCFMLPANYAFIRYCTRGGVPVPTGPTSLITWKFTIQTVPWGLVFLGGGFAL 417  
 DB 369 AIFISLVNFIIPSKIPGLMOPDKPKGLKA--PPAILTWKTVNDKMPNIVILLGGGFAL 426  
 QY 418 AEGSKQSGMAKLGNALIGLKVLPNSVLLVAVELTAFSSNAVAIANIIPVLAEMS 477  
 DB 427 AKGSEQSGLEWGLDKLTPLQHPSPSATAVILCLLTAIFTECTSNVATTLFPLILASWA 486  
 QY 478 LAIEIHPYLIILPAGLACSMFHLVSTPPNALVAGYANIRTKDMAIAGIGPTIITITL 537  
 DB 487 QAICLHPYVYMLPCTLAASLAFMLPVATPPNNAIVFSGGLKVDMARAGFLNIIIGVLAI 546  
 QY 538 FVFCOTWGLVYVYNLNSFPPEWAIYAAAALGNKTH 572  
 DB 547 TLSINSWSIPIF-KLDTFFPSWAHNSNTSQCLLNPSN 580

RESULT 6

US-10-167-994-12

; Sequence 12, Application US/10167994

; Publication No. US20030082647A1

; GENERAL INFORMATION:

; APPLICANT: Reenan, Robert A.

; APPLICANT: Rogina, Blanka

; APPLICANT: Helfand, Stephen L.

; TITLE OF INVENTION: TRANSPORTER PROTEIN

; FILE REFERENCE: 13407-013001

; CURRENT APPLICATION NUMBER: US/10/167,994

; CURRENT FILING DATE: 2002-06-12

; PRIOR APPLICATION NUMBER: US 60/255,013

; PRIOR FILING DATE: 2000-12-12

; PRIOR APPLICATION NUMBER: US 10/017,479

; PRIOR FILING DATE: 2001-12-12

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 587

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-10-167-994-12

Query Match 34.8%; Score 1042.5; DB 14; Length 587;

Best Local Similarity 38.1%; Pred. No. 5.9e-93;

Matches 219; Conservative 129; Mismatches 190; Indels 37; Gaps 8;

QY 26 LVVFLVLLCLPVMNLNEGAFRCMYLLVMAIFWVTEALPLYVTSMIPVAPFIMGIMS 85  
 DB 15 LIVLCLPIFLPLPLIVQTKYACAYSIILMALLWCTEALPLAVTALFVILFPLMGIND 74  
 QY 86 SDQTCRLYFKDTLVFMWGGIMVALAVEYCNLHKLRLALRVQIVGCSRRLHFGILVMTWF 145  
 DB 75 ASEVCIEYFKDTNLFVGGIMVALAVEYCNLHKLRLALRVQIVGCSRRLHFGILVMTWF 134  
 QY 146 LSMWISNACTAMMCPITQAVLELQA--QGVCKINHEPOYQIV-----GGNK 191  
 DB 135 LSMWISNTATTAMVPIGHAVLEQLQSGKKDVEGNNNPTFELQEECPQKEVTKLDNGQP 194  
 QY 192 KNEDEP-----PYPTKITLCYVLGIAYASSIGCGGTIIGTATNLTFRGIYEAREFK 242  
 DB 195 VSAPSEPRTKTOEHRHFSQGLSLC-----ICYSASIGGIATLTGTPNLVLOGVNSLFP 250  
 QY 243 NSTEQMDPFTFMYSVPSMLVYLLTFVFLQWHFGLWRPKSEAEQVORGEGADVAKK 302  
 DB 251 QNGNVNFAWFGFAFTMIILLLLAWLQVLFLGVNFRKNFGFGEGERKQA--AFQ 308  
 QY 303 VIDORYKDLGPMSTHEIQVMILFIFWVYFTRKPGIFLGWADLLNSKDIRNSMPT---- 358  
 DB 309 VIKTQYRLGLGPMSPAEKTVTVLVLLVLFVTFREPFGFPFGWGTTFVANEKGQSMASDGTV 368  
 QY 359 -IFVVMCFMLPANYAFIRYCTRGGVPVPTGPTSLITWKFTIQTVPWGLVFLGGGFAL 417  
 DB 369 AIFISLVNFIIPSKIPGLMOPDKPKGLKA--PPAILTWKTVNDKMPNIVILLGGGFAL 426  
 QY 418 AEGSKQSGMAKLGNALIGLKVLPNSVLLVAVELTAFSSNAVAIANIIPVLAEMS 477  
 DB 427 AKGSEQSGLEWGLDKLTPLQHPSPSATAVILCLLTAIFTECTSNVATTLFPLILASWA 486  
 QY 478 LAIEIHPYLIILPAGLACSMFHLVSTPPNALVAGYANIRTKDMAIAGIGPTIITITL 537  
 DB 487 QAICLHPYVYMLPCTLAASLAFMLPVATPPNNAIVFSGGLKVDMARAGFLNIIIGVLAI 546  
 QY 538 FVFCOTWGLVYVYNLNSFPPEWAIYAAAALGNKTH 572  
 DB 547 TLSINSWSIPIF-KLDTFFPSWAHNSNTSQCLLNPSN 580

RESULT 7

US-10-718-359-13

; Sequence 13, Application US/10718359

; Publication No. US20050095240A1

; GENERAL INFORMATION:

; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.

; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION

; FILE REFERENCE: 275.0008 0101

```

; CURRENT APPLICATION NUMBER: US/10/718,359
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/428,469
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 60/459,441
; PRIOR FILING DATE: 2003-04-01
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 587
; TYPE: PRT
; ORGANISM: rat Nadc1
US-10-718-359-13

```

[illegible]

RESULT 8  
US-10-718-359-12  
; Sequence 12, Application US/10718359  
; Publication No. US20050095240A1  
; GENERAL INFORMATION:  
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.  
; TITLE OF INVENTION: NaCT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION  
; FILE REFERENCE: 275-0008 0101  
; CURRENT APPLICATION NUMBER: US/10/718,359  
; CURRENT FILING DATE: 2003-11-20  
; PRIOR APPLICATION NUMBER: 60/428,469  
; PRIOR FILING DATE: 2002-11-22  
; PRIOR APPLICATION NUMBER: 60/459,441  
; PRIOR FILING DATE: 2003-04-01

```

; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 12
; LENGTH: 581
; TYPE: PRT
; ORGANISM: zebrafish NaCT
US-10-718-359-12

Query Match      34.7%; Score 1039; DB 17; Length 581;
Best Local Similarity 38.0%; Pred. No. 1.3e-92;
Matches 221; Conservative 117; Mismatches 185; Indels 58; Gaps 10

Qy 23 WK----GLVFLVPLLCPLPVMLLNAGAPRCMYLLLVMAIFVWTEALPLYVTSMIDIVAPP 799
Db 10 WKMKNTLLIFCTPFLLLPLPLVIGSGEAGCAVYVVLMAVYVCTEVLPLAVTLPAVLPLFP 69
Qy 80 IMGIMSSDQTCRLYFKDTLVMFMFGIMVALAVECNLHKRIALRLVQIVGCSPRLHGL 139
Db 70 LFRIMESQDVCMYLKDNTNMLFLGLGMVAVAVEHNLHKRIALRLVLLVGVRPALLMLGF 129
Qy 140 IMVTMFLSMISNAACATAMCPITIAVLEELOAGQGVCKINHEPOQIVGVGNKNNEDEPP 199
Db 130 MGVTAFLSMISNATATTAMVPIVQAVLEQLN----NTAQEGSSIFETEEKSTEKQPE 184
Qy 200 YPTK-----ITLCYYLGIAYASSLGGCGTITIGT 227
Db 185 SPGEKVVVLNGDNFMSDEPPEHSREABERLMSKGLTLC----VCYAASIGGTATLTGT 240
Qy 228 ATNLTFFGIYEAREPKNSTEQWDEFTFMFYSPVSMIVYTLTFVFLQWFMGLWRPKSKEA 287
Db 241 GPNLVLMGQMSQLFPDNDPIINFASFVGFAPFNMIIMLTAWLMLQIVFLGFINFKKTWGC 300
Qy 288 QEVORREGADVAKKIDQRYKDLGPMISHIIOVMILFIEMVMYFTRKPGIFLGGWADLL 347
Db 301 GTVYKEEIA--AYNVIKEERSLGPMTFGLSLVALFILLVLVWFTDPGFVDGWATRP 358
Qy 348 NSKD---IRNSMPTIFVVMCMFLPANYAFURYC---TRRGGPVP---TGTPPSLIITWKF 398
Db 359 FNADKEFVTDATVAVFVAALLFVPSKPP--RLCFWRITESFDTVFQSSGPTALLTWKV 416
Qy 399 IQTKVPGLVFLGGGFALAEAGSKSQGWAKLIGNALGLKVLPSVLLLVILVAVFLTA 458
Db 417 TQKMPMSVITLLGGGFALAKGSIISGLSKWLGDQMPLOQIPPMWIAIVTCLMTATFTE 476
Qy 459 FSSNVALANTIIIPVLASMLAIEHTPLYLILPAGLACSMAPHLPVSTPPNALVAGYANIR 518
Db 477 CTNSVATATFLFPLASKMSOSIGVNPPLYVMVPCTLSASFAPMLPVATPPNNAIVSYGILK 536
Qy 519 TKDMAIAGIPTITITITLTFVFCQTGVLGVYVNLNSPFEWA 559
Db 537 VSDMAKTGIWNNIIGHSITLAINWSGRAIF-SLDTFFESWA 576

```

```

RESULT 9
US-10-167-994-3
; Sequence 3, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 592

```

; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-167-994-3

Query Match 34.3%; Score 1026; DB 14; Length 592;  
 Best Local Similarity 37.2%; Pred. No. 2.5e-91;  
 Matches 219; Conservative 124; Mismatches 189; Indels 56; Gaps 10;  
 QY 26 LVVFLVPLLCPLPVMLLNEGAFRCMYLLVMAIFWVTEALPLVYTSMPPIVAFPIGIMS 85  
 DB 15 LIVFFVPIILLPLPLVPSKEAYCAVAILMALFWCTEALPLAVTALFPLILFPMGIVD 74  
 QY 86 SDQTCRLYKFDTLVFMFGIMVALAVEYCNLHKRLALRVIOVGCSPRRLHFLGIMVTF 145  
 DB 75 ASEVAVEYLKDSNLLFFGGLLVAVAEVHNLHKLALRVLLVIGVRAPLILGFMVLTAF 134  
 QY 146 LSMWISNAACTAMCPIIOAVLELQAQGVCKINHEPOYQIVGKNKKNDEDEPPYTK-- 203  
 DB 135 LSMWISNTATSAMVPVIAHVDLQHSQASSNVEE-----GSNNPTFELQEPSQKEV 188  
 QY 204 -----ITLCYVLGIAYASSLGGCGTIIIGTATNLTFTKG 235  
 DB 189 TKLDNGQALPVTSASSEGRAHLSQKHLHLTQCMSLCVCYSASIGGIATLTGTAPNLVLQ 248  
 QY 236 IYEARFNKSTEQMDPTFPMFYSVPSMLVYLLTFVFLQHFHMLG-WRPKSKAEQVQRGR 294  
 DB 249 QINSIFPQNGVNVFASFWSFAPFTWVILLLLAWLQILFLGFNFRKNFGIGEMQ--- 305  
 QY 295 EGADVAKKVIDQRYKDLGPMHSIHIQVMIILFIFMVVMYFTRKPGIFLGWADLL--NSKD- 351  
 DB 306 EQQAAYCVIQTREHLLGPMWFAEKASISILFVLVLLWFTREPGFLGNGNLAPNNAKE 365  
 QY 352 --IRNSMPTIFVVMCFMPLPANYAFLYRYCTRRGGPV--PTGPTPSLITWKFIQTKVPWGL 407  
 DB 366 SWSDGTVAIFIGIIMFIIPSKFGLTQDPENPKLAPLG---LDDWKTVMQKMPWNI 421  
 QY 408 VFLGGGFALAGSKQSGMAKLGKLVLPNSVLLVLLVILVAVLTAFFSSNVAIAN 467  
 DB 422 VLLGGYALAGSKERSGLSEWLNKLTPLQSVPAIPAIAIILSLVATFTECTSNVATT 481  
 QY 468 IIPVLAEMSLAIEIHPYLLIPLAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGI 527  
 DB 482 IFLPILASMAQAICLHPLVYMLPCTLATSLAFMLPVATPPNAIVFSPGDLKVLDMARAGF 541  
 QY 528 GPTIITITLTVFCQWGLVVPVNLGSPPEWQAIIYAAA---ALGNKT 571  
 DB 542 LLNIIGVLIILAINSWGIPLF-SLHSPFSPAQSNTTAQCLPSLANTT 588

RESULT 10  
 US-10-167-994-11  
 ; Sequence 11, Application US/10167994  
 ; Publication No. US20030082647A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Keenan, Robert A.  
 ; APPLICANT: Rogina, Blanka  
 ; APPLICANT: Helfand, Stephen L.  
 ; TITLE OF INVENTION: TRANSPORTER PROTEIN  
 ; FILE REFERENCE: 13407-013001  
 ; CURRENT APPLICATION NUMBER: US/10/167,994  
 ; CURRENT FILING DATE: 2002-06-12  
 ; PRIOR APPLICATION NUMBER: US 60/255,013  
 ; PRIOR FILING DATE: 2000-12-12  
 ; PRIOR APPLICATION NUMBER: US 10/017,479  
 ; PRIOR FILING DATE: 2001-12-12  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 11  
 ; LENGTH: 592  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-167-994-11

Query Match 34.3%; Score 1026; DB 14; Length 592;  
 Best Local Similarity 37.2%; Pred. No. 2.5e-91;  
 Matches 219; Conservative 124; Mismatches 189; Indels 56; Gaps 10;  
 QY 26 LVVFLVPLLCPLPVMLLNEGAFRCMYLLVMAIFWVTEALPLVYTSMPPIVAFPIGIMS 85  
 DB 15 LIVFFVPIILLPLPLVPSKEAYCAVAILMALFWCTEALPLAVTALFPLILFPMGIVD 74  
 QY 86 SDQTCRLYKFDTLVFMFGIMVALAVEYCNLHKRLALRVIOVGCSPRRLHFLGIMVTF 145  
 DB 75 ASEVAVEYLKDSNLLFFGGLLVAVAEVHNLHKLALRVLLVIGVRAPLILGFMVLTAF 134  
 QY 146 LSMWISNAACTAMCPIIOAVLELQAQGVCKINHEPOYQIVGKNKKNDEDEPPYTK-- 203  
 DB 135 LSMWISNTATSAMVPVIAHVDLQHSQASSNVEE-----GSNNPTFELQEPSQKEV 188  
 QY 204 -----ITLCYVLGIAYASSLGGCGTIIIGTATNLTFTKG 235  
 DB 189 TKLDNGQALPVTSASSEGRAHLSQKHLHLTQCMSLCVCYSASIGGIATLTGTAPNLVLQ 248  
 QY 236 IYEARFNKSTEQMDPTFPMFYSVPSMLVYLLTFVFLQHFHMLG-WRPKSKAEQVQRGR 294  
 DB 249 QINSIFPQNGVNVFASFWSFAPFTWVILLLLAWLQILFLGFNFRKNFGIGEMQ--- 305  
 QY 295 EGADVAKKVIDQRYKDLGPMHSIHIQVMIILFIFMVVMYFTRKPGIFLGWADLL--NSKD- 351  
 DB 306 EQQAAYCVIQTREHLLGPMWFAEKASISILFVLVLLWFTREPGFLGNGNLAPNNAKE 365  
 QY 352 --IRNSMPTIFVVMCFMPLPANYAFLYRYCTRRGGPV--PTGPTPSLITWKFIQTKVPWGL 407  
 DB 366 SWSDGTVAIFIGIIMFIIPSKFGLTQDPENPKLAPLG---LDDWKTVMQKMPWNI 421  
 QY 408 VFLGGGFALAGSKQSGMAKLGKLVLPNSVLLVLLVILVAVLTAFFSSNVAIAN 467  
 DB 422 VLLGGYALAGSKERSGLSEWLNKLTPLQSVPAIPAIAIILSLVATFTECTSNVATT 481  
 QY 468 IIPVLAEMSLAIEIHPYLLIPLAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGI 527  
 DB 482 IFLPILASMAQAICLHPLVYMLPCTLATSLAFMLPVATPPNAIVFSPGDLKVLDMARAGF 541  
 QY 528 GPTIITITLTVFCQWGLVVPVNLGSPPEWQAIIYAAA---ALGNKT 571  
 DB 542 LLNIIGVLIILAINSWGIPLF-SLHSPFSPAQSNTTAQCLPSLANTT 588

RESULT 11  
 US-09-729-094-4  
 ; Sequence 4, Application US/09729094  
 ; Patent No. US20020019028A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHATURVEDI, Kabir et al  
 ; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
 ; TITLE OF INVENTION: AND USES THEREOF  
 ; FILE REFERENCE: CL000662  
 ; CURRENT APPLICATION NUMBER: US/09/729,094  
 ; CURRENT FILING DATE: 2000-12-05  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 619  
 ; TYPE: PRT  
 ; ORGANISM: Xenopus laevis  
 US-09-729-094-4

Query Match 33.8%; Score 1010.5; DB 9; Length 619;  
 Best Local Similarity 35.3%; Pred. No. 8.8e-90;  
 Matches 218; Conservative 120; Mismatches 192; Indels 87; Gaps 7;  
 QY 17 NFFANHWKGLVVFVPLLCPLPVMLLNEGAFRCMYLLVMAIFWVTEALPLVYTSMPV 76  
 DB 13 NYF-----IIFLPLVPLLLPLVVPVTKASCGFVIIIMALFWCTEALPLAVTALFPVL 65



Db 127 GFMGTALLSMWISNTATTAMVPIVEAILQOMBEATSAEAGLELVDKGAKEPLGSGV 186  
 QY 187 VGNKKNNEDEPPYTKITLCYILGIAYASSLGGGIIIGTATNLTKGIYEAFKFNSTE 246  
 Db 187 IFEGTLLGQOEDQERKRLCKAMTLCICVAAISIGGTATLTGPNVLLGQNMELFPDSKD 246  
 QY 247 QMDFTFMFYSVPSMLVYTLTLTFVFLQWHEM-----GLWRPKSKEAQEVQREG 296  
 Db 247 LVNFASFASFAPNMLVLLFAWLWLFQVYMFNFKKSGCGLESKNEKA----- 297  
 QY 297 ADVAKVIDQRYKOLGPMGSIHEIQVMILFIMVVMYFTRKPGIFLGWADLL-----NSKDI 352  
 Db 298 ---ALKVLOEYRKLGLPLSFAEINVLIICFFLLVILWFSRDPGFMGLTVAWVEGETKYV 354  
 QY 353 RNSMPTIFVVMCMLPA---NYAFLRYCTRRGGPVPTGTPPSLITWKFOTKVPWGLVF 409  
 Db 355 SDATVAIFVATLLFIVPSQKPFNF-RSQTEBERKTPFPY-PPLDDWKVTOEKVPWGLV 412  
 QY 410 LLGGGFALAEKSGKQSMKALIGNALIGLVKLPNSVLLLV-ILVAVFLTAFSSNVAIANI 468  
 Db 413 LLGGGFALAKGSEASGLSVWGMKQMEPLHVPAAITLILSLLVAVF-TECTSNVATTL 471  
 QY 469 IIPVLAEMSLAIEIHPYLIILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIG 528  
 Db 472 FLPIFASMSRSIGLNPLIYMLPCTLSASFAPMLPVPATPPNAIVFTYGHKLKVDWVKTVI 531  
 QY 529 PTIITITLTFVFCOTWGLVVPNLNSPPEWAOI 561  
 Db 532 MNIIIGVFCVFLAVNTWGRAIF-DLDHFPDWMNV 563

RESULT 14  
 US-10-403-161-72  
 ; Sequence 72, Application US/10403161  
 ; Publication No. US2004004930A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Anderson, David et al.  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: 21402-573C  
 ; CURRENT APPLICATION NUMBER: US/10/403,161  
 ; CURRENT FILING DATE: 2003-03-31  
 ; PRIOR APPLICATION NUMBER: 60/370349  
 ; PRIOR FILING DATE: 2002-04-05  
 ; PRIOR APPLICATION NUMBER: 60/384543  
 ; PRIOR FILING DATE: 2002-05-30  
 ; PRIOR APPLICATION NUMBER: 60/370969  
 ; PRIOR FILING DATE: 2002-04-08  
 ; PRIOR APPLICATION NUMBER: 60/403748  
 ; PRIOR FILING DATE: 2002-08-15  
 ; PRIOR APPLICATION NUMBER: 60/372019  
 ; PRIOR FILING DATE: 2002-04-12  
 ; PRIOR APPLICATION NUMBER: 60/374379  
 ; PRIOR FILING DATE: 2002-04-22  
 ; PRIOR APPLICATION NUMBER: 09/779679  
 ; PRIOR FILING DATE: 2001-02-08  
 ; PRIOR APPLICATION NUMBER: 60/181045  
 ; PRIOR FILING DATE: 2000-02-08  
 ; PRIOR APPLICATION NUMBER: 10/055877  
 ; PRIOR FILING DATE: 2002-01-22  
 ; PRIOR APPLICATION NUMBER: 60/262892  
 ; PRIOR FILING DATE: 2001-01-19  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 173  
 ; SOFTWARE: Curaseq1 version 0.1  
 ; SEQ ID NO 72  
 ; LENGTH: 568  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-403-161-72

Query Match 33.5%; Score 1004; DB 15; Length 568;  
 Best Local Similarity 37.2%; Pred. No. 3.4e-89;  
 Matches 213; Conservative 121; Mismatches 193; Indels 46; Gaps 11;

QY 19 FANHKKGLVV-FLVPLLCIPVMLINEGAEFRCMYLLVMALFWVTEALPLYTSMIPIVA 77  
 Db 7 YVSKFKSFVILFVTPLLLPLVILPAKFRVCAYIILMAIYCTEVIPLAVTSLMEVLL 66  
 QY 78 FPIGMIMSDDQTRLYFKDTLVNFMGGMVAVALEYCNLHKRLALRVIOIVGCSPPRLHF 137  
 Db 67 FPLFQILDSRQVCVQMKOTNMLFGLGLIVAVAVERNLHKRIALRTILWVGAKPARLML 126  
 QY 138 GLIMVTMFSMWTISNAACATMMCPITIOAVLEELQAQGV-----KINHEPOYQI 186  
 Db 127 GFMGTALLSMWISNTATTAMVPIVEAILQOMBEATSAEAGLELVDKGAKEPLGSGV 186  
 QY 187 VGNKKNNEDEPPYTKITLCYILGIAYASSLGGGIIIGTATNLTKGIYEAFKFNSTE 246  
 Db 187 IFEGTLLGQOEDQERKRLCKAMTLCICVAAISIGGTATLTGPNVLLGQNMELFPDSKD 246  
 QY 247 QMDFTFMFYSVPSMLVYTLTLTFVFLQWHEM-----GLWRPKSKEAQEVQREG 296  
 Db 247 LVNFASFASFAPNMLVLLFAWLWLFQVYMFNFKKSGCGLESKNEKA----- 297  
 QY 297 ADVAKVIDQRYKOLGPMGSIHEIQVMILFIMVVMYFTRKPGIFLGWADLL-----NSKDI 352  
 Db 298 ---ALKVLOEYRKLGLPLSFAEINVLIICFFLLVILWFSRDPGFMGLTVAWVEGETKYV 354  
 QY 353 RNSMPTIFVVMCMLPA---NYAFLRYCTRRGGPVPTGTPPSLITWKFOTKVPWGLVF 409  
 Db 355 SDATVAIFVATLLFIVPSQKPFNF-RSQTEBERKTPFPY-PPLDDWKVTOEKVPWGLV 412  
 QY 410 LLGGGFALAEKSGKQSMKALIGNALIGLVKLPNSVLLLV-ILVAVFLTAFSSNVAIANI 468  
 Db 413 LLGGGFALAKGSEASGLSVWGMKQMEPLHVPAAITLILSLLVAVF-TECTSNVATTL 471  
 QY 469 IIPVLAEMSLAIEIHPYLIILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIG 528  
 Db 472 FLPIFASMSRSIGLNPLIYMLPCTLSASFAPMLPVPATPPNAIVFTYGHKLKVDWVKTVI 531  
 QY 529 PTIITITLTFVFCOTWGLVVPNLNSPPEWAOI 561  
 Db 532 MNIIIGVFCVFLAVNTWGRAIF-DLDHFPDWMNV 563

RESULT 15  
 US-10-490-080-1  
 ; Sequence 1, Application US/10490080  
 ; Publication No. US20040253597A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Takeda Chemical Industries, Ltd.  
 ; TITLE OF INVENTION: Novel Protein and its DNA  
 ; FILE REFERENCE: P02-0109PCT  
 ; CURRENT APPLICATION NUMBER: US/10/490,080  
 ; CURRENT FILING DATE: 2004-03-17  
 ; PRIOR APPLICATION NUMBER: JP 2001-281992  
 ; PRIOR FILING DATE: 2001-09-17  
 ; PRIOR APPLICATION NUMBER: JP 2001-306873  
 ; PRIOR FILING DATE: 2001-10-02  
 ; PRIOR APPLICATION NUMBER: JP 2002-113279  
 ; PRIOR FILING DATE: 2002-04-16  
 ; NUMBER OF SEQ ID NOS: 42  
 ; SEQ ID NO 1  
 ; LENGTH: 568  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-10-490-080-1

Query Match 33.5%; Score 1004; DB 16; Length 568;  
 Best Local Similarity 37.2%; Pred. No. 3.4e-89;  
 Matches 213; Conservative 121; Mismatches 193; Indels 46; Gaps 11;

QY 19 FANHKKGLVV-FLVPLLCIPVMLINEGAEFRCMYLLVMALFWVTEALPLYTSMIPIVA 77  
 Db 7 YVSKFKSFVILFVTPLLLPLVILPAKFRVCAYIILMAIYCTEVIPLAVTSLMEVLL 66



**THIS PAGE IS BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:29:45 ; Search time 17.0834 Seconds  
(without alignments)  
3221.603 Million cell updates/sec

Title: US-10-017-479A-2

Perfect score: 2994

Sequence: 1 MEIEIGEPQPPVKCSNFFA.....NSFPWAQIYAAALGNKTH 572

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR 79:\*

1: piri:\*

2: piri2:\*

3: piri3:\*

4: piri4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1024	34.2	593	2	146528 sodium/dicarboxyla
2	1003	33.5	568	2	JC7911 Na+/coupled citrat
3	885	29.6	595	2	A47714 Na+/sulfate cotran
4	859	28.7	599	2	T21613 hypothetical prote
5	815	27.2	545	2	T18694 hypothetical prote
6	773	25.8	526	2	S43561 YCR37C homolog K08
7	773	25.8	534	2	G88575 protein K08E5.2 (i
8	759.5	25.4	539	2	A88546 protein R107.1 (im
9	730.5	24.4	539	2	S30871 hypothetical prote
10	676.5	22.6	520	2	C89980 hypothetical prote
11	632	21.1	552	2	F64546 sodium-dependent t
12	626	20.9	548	2	E71961 probable transpor
13	511	17.1	446	2	F70302 transporter (pho8
14	447	14.9	461	2	I64080 probable membr
15	445.5	14.9	471	2	A81948 probable transmemb
16	444.5	14.8	471	2	G81156 transporter, Nadc
17	440.5	14.7	443	2	C69205 sodium/dicarboxyla
18	426.5	14.2	462	2	B82510 transporter, Nadc
19	411	13.7	456	2	AF0093 Sodium, sulfate sym
20	391.5	13.1	432	2	H64383 Na+ transporter -
21	393	12.8	487	2	C82215 probable transpor
22	374	12.5	867	2	T40336 probable MSF trans
23	367	12.3	478	2	A82336 probable transpor
24	336.5	11.2	517	2	C89840 conserved hypothet
25	302.5	10.1	894	2	S45135 probable membr
26	300.5	10.0	513	2	I39534 hypothetical prote
27	281.5	9.4	450	2	A10909 probable membr
28	277	9.3	881	2	S46633 probable membr
29	244.5	8.2	923	1	MMBY7C probable membr

30	212	7.1	425	2	C75019 transport protein
31	209.5	7.0	424	2	F71205 hypothetical prote
32	190	6.3	157	2	S36784 mucin - rat (fragm
33	188.5	6.3	589	2	H84072 sodium/sulfate sym
34	185	6.2	482	2	F71969 hypothetical prote
35	184	6.1	479	1	B64043 conserved hypothet
36	184	6.1	487	1	Q8ECRS YGJE protein - Esc
37	179	6.0	618	2	F84409 arsenite transport
38	177.5	5.9	589	2	F75398 hypothetical prote
39	177	5.9	487	2	B91122 hypothetical prote
40	177	5.9	487	2	A85967 sulfur deprivation
41	176.5	5.9	612	2	S74936 hypothetical prote
42	164.5	5.5	592	2	D53393 hypothetical prote
43	164	5.5	470	2	G86516 dicarboxylase tran
44	162	5.4	470	2	H72105 arsenical pump mem
45	161	5.4	429	1	B41902 gene DN10 protein
46	160.5	5.4	838	2	S28911 hypothetical prote
47	159	5.3	449	2	S76839 citrate carrier (i
48	159	5.3	501	2	A50578 probable membr
49	158	5.3	475	2	A10010 probable dicarboxy
50	155	5.2	471	2	F71543 arsenite transport
51	152.5	5.1	411	2	G90154 ocucutaneous alb
52	152.5	5.1	532	2	A57173 probable integral
53	150.5	5.0	469	2	T35526 hypothetical prote
54	150	5.0	428	2	G83803 arsenical pump mem
55	150	5.0	429	1	C41903 hypothetical prote
56	148	4.9	430	2	G89962 probable ion trans
57	148	4.9	610	2	AG0312 ybhl protein - Esc
58	146	4.9	477	1	B64813 probable integral
59	146	4.9	488	2	B81831 C4-dicarboxylate t
60	145	4.8	488	2	A81194 probable sodium/au
61	145	4.8	608	2	AC0798 conserved hypothet
62	143	4.8	425	2	F72315 probable membr
63	142	4.7	477	2	G85579 probable membr
64	142	4.7	477	2	F90728 ybds protein - Esc
65	142	4.7	487	1	B84795 probable membr
66	142	4.7	487	2	G85560 probable membr
67	142	4.7	487	2	C90710 conserved hypothet
68	140.5	4.7	436	2	D98100 sulfur deprivat
69	140	4.7	585	2	S69216 2-oxoglutarate/mal
70	136.5	4.6	478	1	F69811 probable membr
71	136.5	4.6	484	2	A50887 probable sodium/au
72	135	4.5	610	2	H83165 membrane protein (
73	134.5	4.5	436	2	C95236 transporter, sodiu
74	134	4.5	590	2	AG3106 AcrB/AcrD/AcrF fam
75	134	4.5	590	2	E98180 transporter, sodiu
76	133.5	4.5	1046	2	AD2959 probable rnd efflu
77	133.5	4.5	1046	2	B98324 probable transpor
78	133	4.4	610	2	A85870 probable transpor
79	133	4.4	610	2	H91025 probable transpor
80	133	4.4	610	2	B65001 conserved hypothet
81	132	4.4	421	2	G72300 probable arsB prot
82	131.5	4.4	428	2	H70528 sulfur deprivation
83	131	4.4	620	2	H82761 CAMP-activated Na+
84	129.5	4.3	759	2	A46188 hypothetical prote
85	128.5	4.3	461	2	A82185 conserved hypothet
86	127	4.2	489	2	AF2970 hypothetical prote
87	127	4.2	489	2	D98312 gluconate transpor
88	124.5	4.2	447	2	C86130 gluconate transpor
89	124.5	4.2	447	2	H91288 gluconate permease
90	124.5	4.2	447	2	S56546 arsenic efflux pum
91	120	4.0	431	2	A11599 hypothetical prote
92	119	4.0	440	2	G84069 probable arsenical
93	118	3.9	428	2	E81324 sulfur deprivation
94	118	3.9	548	2	AC3587 probable integral
95	116	3.9	430	2	C81402 NADH2 dehydrogenas
96	116	3.9	495	1	S26870 involved in spore
97	115	3.8	522	2	A83804 2-oxoglutarate/mal
98	115	3.8	527	2	T49900 probable C4-dicarb
99	114.5	3.8	427	2	B83001 hypothetical prote
100	114	3.8	372	2	G90740

Thu Jun 30 13:13:30 2005

ALIGNMENTS

```

RESULT 1
146528
sodium/dicarboxylate cotransporter - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 09-Jul-2004
C:Accession: I46528
R:Paigor, A.M.
J. Biol. Chem. 270, 5779-5785, 1995
A:Title: Sequence and functional characterization of a renal sodium/dicarboxylate cotran
A:Reference number: I46528; MUID:95197598; PMID:7890707
A:Accession: I46528
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-593 <PAJ>
A:Cross-references: UNIPROT:Q28615; EMBL:U12186; NID:G758383; PIDN:AAA99666.1; PID:G7583
C:Superfamily: sodium/sulfate cotransporter

Query Match 34.2%; Score 1024; DB 2; Length 593;
Best Local Similarity 37.6%; Pred. No. 5.1e-77;
Matches 219; Conservative 127; Mismatches 186; Indels 50; Gaps 10;

QY 20 ANHWKG-----LVVFLVLLCLPVMMLNEGAEFRGMYLLVMAIFWVTEALPLYVTSM 72
DB 2 ATCQGLWAYRMYLLVFLPLSLPLVPRKEAYCAYALLMALFWCTDALPLVATL 61

QY 73 IPIVAFPIGIMSSDQTCRLYFKDITLVFMFGGIMVALAVEYCNLHKRLALRVLIQVGCSP 132
DB 62 LPLCLFPWKGIMEASEVGLYKDTNVLFIGLLALAVEHNLHKRIALRVLLITGVPR 121

QY 133 RLHFLGLIMVTFMSMISNAACATMPCPIIQAIVLELO-----AQGVCKINHE----- 181
DB 122 ALLILGFVMTAFSLMISNTASTAMVPIAHAVLQELNLTQSNVEEGSDNPTELOEPS 181

QY 182 PQYQIVGGKKNNEDEPPYPT-----KITLCVYLGIAYASSLGCGT 223
DB 182 POKETSQVDEKONGQAPLPAVPLESGEHTVQBLRFSQNSLC-----VCYSASIGGIAT 237

QY 224 IIGTATNLTFKGIYEAREPKNSTEQMDFTFMFVSPSLVYTLTFFVFLQWHFGLWRPK 283
DB 238 LTGTPNLVLQOMTSLFPQNPVNFASWFGFAPIPMVILLLSWLMLQLFLGIFNPK 297

QY 284 SKBAQEVORGBADVAKVVDQYKDLGPMSEIHIQWMLFIFWVMYFTRKPGIELGW 343
DB 298 NFGIREQEHQORQAAYRVVITQYRLGLGPMSEFAKAVFLLFVILVLLWFTREPGFHG 357

QY 344 ADLNSKD-----IRNSMPTIFVVMCFMLPANYAFRYCTRGGPVPTGPTPSLITWKF 398
DB 358 GNLVFSASGRVWSDGSASILIGVLFPMVPSKIPGLTQDPNPGRLKA--PPALLNWK 415

QY 399 IQTKVPWGLVFLGGGFALAEKSGKQMAKLIQNALIGLK-VLPNSVLLVILVAVFLT 457
DB 416 VNKMPWNLVLLGGYALAKGSESGLSQWLGNKMLPQHVPPATVFIICLVATF-T 474

QY 458 AFSSNVATANIILPVAENSLAIEIHPYLLIPLAGLACSMAFHLPVSTPPNALVAGYANI 517
DB 475 ECTSNAAATLTLPLILASQAQICLHPLVMDPCTLASLAFMLPVATPPNAIVFSFGL 534

QY 518 RYKDMAIAGIGPTIITILFVFCQTWGLVVPNLNSFPEWA 559
DB 535 RYSDMARAGIMLNIIGVLVIMLAINSGVPMF-QLHTFPEWA 575

RESULT 2
JC7911
Na+/sulfate cotransporter NaCT - human
C:Species: Homo sapiens (man)
C>Date: 31-Mar-2003 #sequence_revision 31-Mar-2003 #text_change 14-Jul-2003
C:Accession: JC7911
R:Inoue, K.; Zhuang, L.; Ganapathy, V.
Biochem. Biophys. Res. Commun. 299, 465-471, 2002

```

```

A:Title: Human Na+-coupled citrate transporter: Primary structure, genomic organization,
A:Reference number: JC7911; MUID:22334959; PMID:12445824
A:Accession: JC7911
A:Molecule type: mRNA
A:Residues: 1-568 <INO>
A:Cross-references: GB:AV151833
C:Comment: This transporter classified as a tricarboxylate transporter represents the fi
the cellular entry of citrate by a process energized by the electrochemical Na+ gradient
or the synthesis of fatty acid and chol 17p12-13
C:Genetics:
A:Gene: nact

Query Match 33.5%; Score 1003; DB 2; Length 568;
Best Local Similarity 37.1%; Pred. No. 2.7e-75;
Matches 214; Conservative 124; Mismatches 185; Indels 54; Gaps 13;

QY 19 FANHKGGLVV-FLVPLLCPLPVMMLNEGAEFRGMYLLVMAIFWVTEALPLYVTSMPIVA 77
DB 7 YVSKFKSFVILEFTPLLLPLVILMPAKFVRCAIYIILMAIYMCVEIPLAVTSLMPVLL 66

QY 78 FPIGIMSSDQTCRLYFKDITLVFMFGGIMVALAVEYCNLHKRLALRVLIQVGCSPRRHF 137
DB 67 FFLFQLIDRSQVCQVMKDTNMLFLGGLIVAVAVERNLHKRIALRVLLVWGAKPRLML 126

QY 138 GLIMVTFMSMISNAACATMPCPIIQAIVLELOAGVC-----KINHEPQYQI 186
DB 127 GFMGVTALLSMVISNTATTAMVPIEATLQOMEATSAATEAGLELVLDKGAKEPLGSGV 186

QY 187 VGG-----NKKNEDEPPYPTKITLCVYLGIAYASSLGCGTIIGTATNLTFKGIYEAREK 242
DB 187 IFEGPILGOQDQERKRLCKAMTLC-----ICVAASIGGTATLTGTGPNVVLIGQMNLEFP 242

QY 243 NSTEQMDFTFMFVSPSLVYTLTFFVFLQWHF-----GLMRPKSKEAQEVOR 292
DB 243 DSKDLVNFASWFAFAPFNNMLVLLFPAWLQFVYMRFNFKSWGCGLESKNEKA----- 297

QY 293 GREGADVAKVIDQYKDLGPMSEIHIQWMLFIFWVMYFTRKPGIFIGWADLL-----N 348
DB 298 -----ALKVQBEYRKLGLPLSFABINVLICFFLLVILWFSRDPGFMPLTVANVEGE 350

QY 349 SKDIRNSMPTIFVVMCFMLP-----NYAFRLYCTRGGPVPTGPTPSLITWKFIOTKVPW 405
DB 351 TKTVSDATVAIFVATLFFVPSQKPFNP-RSQTEERKTPFPY-PLLDWKVYQEKVPW 408

QY 406 GLVFLGGGFALAEKSGKQMAKLIQNALIGLKVLPSNVLLVW-ILVAVFLTAFSSNVA 464
DB 409 GIVLLGGGFALAKGSESGLSVMWQMEPLHAVPPAAITLILSLVAVF-TECTSNVA 467

QY 465 IANIILPVAENSLAIEIHPYLLIPLAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAI 524
DB 468 TTTFLFIPFASMSRSIGLNPLYIMLPCTLSASFAPMLPVATPPNALIVFTYGHUKVADWVK 527

QY 525 AGIGPTIITITILFVFCQTWGLVVPNLNSFPEWAQI 561
DB 528 TGVIMNIIGVFCVFLAVNTWGRAIF-DLDHFPDWNV 563

RESULT 3
A47714
Na+/sulfate cotransporter, renal - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004
C:Accession: A47714
R:Markovich, D.; Forgo, J.; Stange, G.; Biber, J.; Murer, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 8073-8077, 1993
A:Title: Expression cloning of rat renal Na+/SO4(2-) cotransport.
A:Reference number: A47714; MUID:93376745; PMID:7690140
A:Accession: A47714
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-595 <MAR>
A:Cross-references: UNIPROT:Q07782; GB:L19102; NID:G310182; PIDN:AAA41677.1; PID:G310183
C:Superfamily: sodium/sulfate cotransporter

```

Query Match 29.6%; Score 885; DB 2; Length 595;  
 Best Local Similarity 34.2%; Pred. No. 1.7e-65;  
 Matches 202; Conservative 123; Mismatches 190; Indels 76; Gaps 14;

QY 26 LVVFLVPLCLPMLNLEGAEPFCMYLLVMAIFWVTEALPLVYVTSWIPVAPINGMS 85  
 DB 15 LVVFTV-LVLLPLPLIIRSKAEACVILFVIATFTEALPISITALLPLGMPFPGMS 73  
 QY 86 SDQTCRLYFKDTLVLMFMGGIMVALAVEYCNLHKRLALRVIOVGCSPRLHFLGIMVTF 145  
 DB 74 STHVASAYFKDFHLLIGVICLATSIEKNLHKRIALRMVMVGVNPAWLTGFSSTAP 133  
 QY 146 LSWISNACTAMMCPDIOAVLEEL-----QAQGVCKINHEPOYQIV 187  
 DB 134 LSWLSNTSTAAMVMPIVEAVAQOITSABAEAEATQMTYFNESAAOGL-----EVDETII 188  
 QY 188 G-GNKKNEDEP-----PYPTK-----ITLCYILGI 212  
 DB 189 GOETNERKEKTRPALGSSNDKGVSSKMETEKNTVTGAKYRSKDHMMCKLMCLC---I 244  
 QY 213 AYASSIGGGCTIIGTATNLTFFGIYKARFNKSTQMDPPTFMFYSVPMSLVTLTFFVL 272  
 DB 245 AYSSITIGGTTTGTSTNLIFSEHFNTRYD-CRCLNFGSWFLFPFPVAVILLLSWML 303  
 QY 273 QNHFMGLMRPKSKEAQEVORREGADVAKKVIDQYKIDGPMSEIHEIQMILFIFMVVY 332  
 DB 304 QMLFLGFNFKMFKCGKTKTLEKA--CAEVIKQEVKLGPMRYQEIIVTLVIFIVMALL 361  
 QY 333 FTRKRGIFLGWADLLNSKD--TRNSNPTTFVVMCMPLPANYAFRYCTRGGPVPTGPT 390  
 DB 362 FSRDFGFTVGSVLSFSEYGYVTDSTVALVAGILFFLIPAKK--LTQMTSTGDIIFDYS 419  
 QY 391 PSLITWKTFTQVPMGLVPLGGLGFPALAGSKOSGMAKLIGNALIGLKLVPNSVLLVVI 450  
 DB 420 P-LITWKEFQSPMPMDIALVGGGFPALAGCQVGLSSWIGSKLPLGSLPWLILISS 478  
 QY 451 LVAVFLTAFSSVAIAIIPVLAEMSLAIEIHPYLLIPAGLACSMAPFLPVSTPPNAL 510  
 DB 479 LIVTSLEVASNPATITLFPILSPALAEIHYVNLPHILLPSTLCTSFAPLLPVANPPNAI 538  
 QY 511 VAGYANIRTKMAIAGIGTITITITLFPVCTQWGLVVP--NLSFPFEWA 559  
 DB 539 VFSYGLHVKVIDMVKAGLVNIGLVAVVMGLMFTW---IEPMENLHEYPSWA 586

RESULT 4  
 T21613  
 Hypochemical protein F31F6.6 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T21613  
 R:Percy, C.  
 submitted to the EMBL Data Library, March 1996  
 A:Reference number: Z19449  
 A:Accession: T21613  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-599 <WIL>  
 A:Cross-references: UNIPROT:Q93455; EMBL:Z69884; PIDN:CAA93752.1; GSPDB:GN00028; CESP:F31F6  
 A:Experimental source: clone F31F6  
 C:Genetics:  
 A:Gene: CESP:F31F6.6  
 A:Map position: X  
 A:Introns: 39/3; 142/2; 207/2; 255/2; 285/2; 397/1; 437/3; 493/2; 537/3  
 C:Superfamily: sodium/sulfate cotransporter

Query Match 28.7%; Score 859; DB 2; Length 599;  
 Best Local Similarity 32.6%; Pred. No. 2.5e-63;  
 Matches 188; Conservative 128; Mismatches 197; Indels 64; Gaps 11;

QY 23 WKGLVFLVPLCLPMLNLE--GAEFRCMYLLVMAIFWVTEALPLVYVTSWIPVAPPI 80

DB 21 WQALLIF-SPLL-----MFVGDHGLQAKCLYCVAVMGSYWVFEALPLAITAIFPMLFPL 75  
 QY 81 MGIMSSDQTCRLYFKDTLVLMFMGGIMVALAVEYCNLHKRLALRVIOVGCSPRLHFLGII 140  
 DB 76 FGIMSESEVARAYLPDTCFLFMGGLMVALAVEKCELHARVALFVLKTVGSEPARVMAGPM 135  
 QY 141 MYTMFLSMWISNACTAMMCPDIOAVLEELQAQ-----GVCKINHEPOYQIVGNGKK- 192  
 DB 136 GVTGFLSMWISNTATTALMWPLQSVITELVSNHRMEDLVALCEAHNHSRKHSGVMRRL 195  
 QY 193 -----NNEDE-----PPYPTKTTLCYILGIAVASSIGGGCTIIGTATNLTFFGIYEA 239  
 DB 196 SLPNENNEIKREEMDTAMSPREQKMAKGLMSVCFSANIGGAATITGTASNVLVYQLNE 255  
 QY 240 RFKSTEQMDPPTFMFYSVPMSLVTLTFFLQWHFMGLMRPKSKEAQEVORREGADV 299  
 DB 256 LPPGADTGVNLSWILFAPF--MVFCCLYICWCVLYLYL-----RDAPKGSII 302  
 QY 300 AKKVIDQYKIDGPMSEIHEIQMILFIFMVVYFTRKPGIFLGWADL----- 346  
 DB 303 VTRKLOQKYNELHAPSPAEMAVIFCFALLVLWLIREPQVVPVGMGEMPKDELVPKSLTEK 362  
 QY 347 ----LNSKDINSMPTIFVVMCMPLPANYAFRYCTRGGPVPTGPTSLITWFIQTK 402  
 DB 363 KNTHTLTPRFVSDATSAFIVILLTLPF-----KLPSRSGSSEQRKASSGLLDWATVQDR 417  
 QY 403 VPWGLVFLGGLGFPALAGSKOSGMAKLIGNALIGLKLVPNSVLLVVLVAVFLTAFSSN 462  
 DB 418 FPWSVFLGGLGFPALAGSKESGLSHDIGAMRYLDVFNHNIIMLICIIISVTLTVNCSN 477  
 QY 463 VAIANITIPVLAEMSLAIEIHPYLLIPAGLACSMAPFLPVSTPPNALVAGYANIRTKDM 522  
 DB 478 TVIASIRIPVLAELARSLEIDPLNFMPLVTTISASFAFLPVATPPNAIVFSSGYLKLVFDM 537  
 QY 523 AIAGIGTITITITLFPVCTQWGLVVPVNLNSFPFEWA 559  
 DB 538 FVSGLCVTLGCVLISMLNMLWAGVFP--NLHLFPQWA 573

RESULT 5  
 T18694  
 Hypochemical protein B0285.6 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
 C:Accession: T18694  
 R:Suleston, J.  
 submitted to the EMBL Data Library, June 1994  
 A:Reference number: Z19007  
 A:Accession: T18694  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-545 <WIL>  
 A:Cross-references: EMBL:Z34533; PIDN:CAA84299.1; GSPDB:GN00021; CESP:B0285.6  
 A:Experimental source: clone B0285  
 C:Genetics:  
 A:Gene: CESP:B0285.6  
 A:Map position: 3  
 A:Introns: 6/3; 56/3; 95/2; 107/2; 160/2; 175/1; 211/2; 287/2; 353/1; 393/3; 494/3  
 C:Superfamily: sodium/sulfate cotransporter

Query Match 27.2%; Score 815; DB 2; Length 545;  
 Best Local Similarity 32.1%; Pred. No. 1e-59;  
 Matches 177; Conservative 119; Mismatches 221; Indels 34; Gaps 8;

QY 43 EGAEFRCMYLLVMAIFWVTEALPLVYVTSWIPVAPIMGIMSSDQTCRLYFKDTLVNFM 102  
 DB 4 ESTEGKCAFVILTMSCYVAEVPVAVTSFIPMIALPFLGIVSIVKEVAPKYFADTNIVFP 63  
 QY 103 GGIMVALAVEYCNLHKRLALRVIOVGCSPRLHFLGIMVTLMSWISNACTAMMCPPI 162  
 DB 64 NSLMSLAVBECQLHKKIALKMLTYVTRPHLWAGMFIITFSLWISDTACCALMAPI 123  
 QY 163 IQAVLEELQAQGVCKINHEPOYQIVGNGKNNEDEPPYPTK-----ITLICY 208

Db 124 AYALLEEIM---IPKMRPEKEHEIEVWKIFDKEDBEEKKKLDTSRLSVDRDRGICKCM 180  
Qy 209 YLGIAVASSLGGCGTIICTATNLTFFKIYEAFKFNSTEQMDFFTFMFYSVPSMLVYTLT 268  
Db 181 MLLVAHASLIGGTGTTNSTGNLIFRDNIEKNFPNEDHGISYLSWMAFAIPPMIFWESS 240  
Qy 269 FVFLOW-----HPMGLWR\_PKSKEAEOVQREGADVAKKVDQRYKDLGPMSEHIOV 321  
Db 241 WFIQLOFLGPRHLGMGFREPTTEKQE-----EEVAKRAVMKSYDQGLGPMTEAKST 293  
Qy 322 MLFIEMVMVYFTRKPGIFLGWADLLNSKDIRNSMPTIFVVMCFMLPANYAFRLVYCTRR 381  
Db 294 LVIFVLAVLSWSDPKVIGNSDLFRKGVYVUSCSGLVAVLLFWPKKKPDR-IFRK 352  
Qy 382 GGPVPTGPTSLITWKFIOQKVPNGLVFLGGGFALAEQSKQSGMAKLIQNALIG-LKVL 440  
Db 353 DKSRPSVRQBELIDWCVRRRFPWSIILLGAGFAISDAVRVSGLSLIACSLNSTISKM 412  
Qy 441 PNSVLLLVILVAVFLTAPSNVAIANIIPVLAEMSLAIEIHPYLLILPAGLACSMAPH 500  
Db 413 PFFVMQIILSVVVVMTFSTNSATSFIPISFKMAEAVGAHPLYFSIPTAIGPSFSFM 472  
Qy 501 LPVSTPPNALVAGYANIRTKDMAGTGPITITITLTFVFCQTWGLVYVYVNLNSFFPEWAQ 560  
Db 473 LPMATPANAIYETKTIEMIDWVSCGVFLNATFCIAITAINMTWAFWFLF-NMGTYPDYAL 531  
Qy 561 IYAAALGNKT 571  
Db 532 RHATNMTGNS 542

RESULT 6

S43561  
YCR37C homolog K08E5.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 07-Sep-1994 #sequence\_revision 10-Nov-1995 #text\_change 24-Nov-1999  
C:Accession: S43561  
R:Kershaw, J.  
submitted to the EMBL Data Library, March 1994  
A:Reference number: S43561  
A:Accession: S43561  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-526 <KER>  
A:Cross-references: EMBL:Z30974  
C:Genetics:  
A:Introns: 33/3; 122/2; 169/2; 214/2; 253/2; 481/3  
C:Superfamily: sodium/sulfate cotransporter

Query Match 25.8%; Score 773; DB 2; Length 526;  
Best Local Similarity 33.2%; Pred. No. 2.9e-56;  
Matches 176; Conservative 106; Mismatches 190; Indels 58; Gaps 9;

Qy 46 EPRCYLLLVMAIFWVTEALPLYVTSMIPVAPPIMSSDOTCRLYFKDTLVNFMGGI 105  
Db 34 EWKCAVCVCIIVYVWSEVMPVAVTAMLPVVLPLVGLVDANTTAKEYMNDTNLFIGGL 93  
Qy 106 MVALAVEYCNLHKRLALRVIQVGCSPRLHFGILMTWFLSMWISNAACACTAMCPIQA 165  
Db 94 IMAAAVEKCDLHERVALSVLRVCGSEPKWIMLGFTVTALLSSFSINTATTAMVPIGQS 153  
Qy 166 VLEELQAQGVCKINHEPQVQIVGGNKKNNEDEPPYPTKITLCYILGIAYASSLGGCGTII 225  
Db 154 VVQQL-----ISSFQHP-----TNGERGLCKKQWATGLVLSICFAANIGGTGTAT 200  
Qy 226 GTATNLTFFKIYEAFKFNSTEQMDFFTFMFYSVPSMLVYTLTTFVFLQHFHGLWRPKSK 285  
Db 201 GTPSNLVMQLGSLALFPKVDGSLNVYTWIFPFAVPLMLLCLFVAVMTLVSFEL-----R 253  
Qy 286 EAOEVQREGADVAKKVIDQRYKDLGPMSEHIOVMILFIWMVYFTRKPGIFLGWAD 345  
Db 254 DAPEKD-----EAVTEMLKTRYNELPRMTYAEKSVFVCFILLSLWFRNPGVVPFGV 307

Qy 346 LLNSKDIRNSMPTIFVVMCFMLP-----ANYAFRLYCTRRGGPVPTGPTPSLITWKF 399  
Db 308 FFKGAYTDATSAMIVAVLLFVLPSERPDLATYIKKEDLKKRG-----CLMDWKTM 358  
Qy 400 QTKVPWGLVFLGGGFALAEQSKQSGMAKLIQNALIGLVLPNSVLLLVILVAVFLTAF 459  
Db 359 QETFPWSVLLGGGFALAAAGVKESGLSLLIGNSLSSIEHLPLWLILQLLTMLIAMVITNI 418  
Qy 460 SNNVAIANIIPVLAEMSLAIEIHPYLLILPAGLACSMAPHLPVSTPPNALVAGYANIR 519  
Db 419 CSNTVTASIFVIVATLAQRAGHPTLMLPTLASSFAIFPVGTGPPNAIVFGSGMWKV 478  
Qy 520 KDMATAGIGPTIITITLTFVFCQTWGLVYVYV-----LNSFFPEWAQI 561  
Db 479 SDMAFVG-----GIISLELLVLT---VLWNYSIAVLTPLLEFPTWAI 519

RESULT 7

G88575  
protein K08E5.2 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C:Accession: G88575  
R:anonymous, The C. elegans Sequencing Consortium.  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_elegans/  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: G88575  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-534 <STO>  
A:Cross-references: GB:chr\_III; PIDN:CAA83225.1; PID:g3878357; GSPDB:GN00021; CESP:K08E5.  
C:Genetics:  
A:Gene: K08E5.2  
A:Map position: 3  
C:Superfamily: sodium/sulfate cotransporter

Query Match 25.8%; Score 773; DB 2; Length 534;  
Best Local Similarity 33.2%; Pred. No. 3e-56;  
Matches 176; Conservative 106; Mismatches 190; Indels 58; Gaps 9;

Qy 46 EPRCYLLLVMAIFWVTEALPLYVTSMIPVAPPIMSSDOTCRLYFKDTLVNFMGGI 105  
Db 42 EWKCAVCVCIIVYVWSEVMPVAVTAMLPVVLPLVGLVDANTTAKEYMNDTNLFIGGL 101  
Qy 106 MVALAVEYCNLHKRLALRVIQVGCSPRLHFGILMTWFLSMWISNAACACTAMCPIQA 165  
Db 102 IMAAAVEKCDLHERVALSVLRVCGSEPKWIMLGFTVTALLSSFSINTATTAMVPIGQS 161  
Qy 166 VLEELQAQGVCKINHEPQVQIVGGNKKNNEDEPPYPTKITLCYILGIAYASSLGGCGTII 225  
Db 162 VVQQL-----ISSFQHP-----TNGERGLCKKQWATGLVLSICFAANIGGTGTAT 208  
Qy 226 GTATNLTFFKIYEAFKFNSTEQMDFFTFMFYSVPSMLVYTLTTFVFLQHFHGLWRPKSK 285  
Db 209 GTPSNLVMQLGSLALFPKVDGSLNVYTWIFPFAVPLMLLCLFVAVMTLVSFEL-----R 261  
Qy 286 EAOEVQREGADVAKKVIDQRYKDLGPMSEHIOVMILFIWMVYFTRKPGIFLGWAD 345  
Db 262 DAPEKD-----EAVTEMLKTRYNELPRMTYAEKSVFVCFILLSLWFRNPGVVPFGV 315  
Qy 346 LLNSKDIRNSMPTIFVVMCFMLP-----ANYAFRLYCTRRGGPVPTGPTPSLITWKF 399  
Db 316 FFKGAYTDATSAMIVAVLLFVLPSERPDLATYIKKEDLKKRG-----CLMDWKTM 366  
Qy 400 QTKVPWGLVFLGGGFALAEQSKQSGMAKLIQNALIGLVLPNSVLLLVILVAVFLTAF 459  
Db 367 QETFPWSVLLGGGFALAAAGVKESGLSLLIGNSLSSIEHLPLWLILQLLTMLIAMVITNI 426  
Qy 460 SNNVAIANIIPVLAEMSLAIEIHPYLLILPAGLACSMAPHLPVSTPPNALVAGYANIR 519

```
Db 427 CSNTVTASIFVPIVATLQAGHHPPTLMLPTTLASSFAPIPVGPPNAIVFGSGMVKV 486
QY 520 KDMAIAGIGPTIITITLFFVFCOTWGLVVPN-----LMSFPPEWAOI 561
Db 487 SDMAFVG-----GIISLELLVLT---VLYMNSIAVLTLPLLEFFPWAI 527

RESULT 8
A88546
protein R107.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: A88546
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: A88546
A>Status: preliminary
A:Molecule type: DNA
A:Cross-references: ENBL:Z14092
A:Residues: 1-539 <STO>
A:Genetic code: SGC4
A:Introns: 30/3; 80/3; 119/2; 217/2; 251/3; 293/2; 390/3; 416/1; 445/3; 491/3
A:Superfamily: sodium/sulfate cotransporter
C:Keywords: transmembrane protein

Query Match 25.4%; Score 759.5; DB 2; Length 539;
Best Local Similarity 31.8%; Pred. No. 46-55;
Matches 177; Conservative 111; Mismatches 205; Indels 63; Gaps 14;

QY 24 KGLVVLPLLLCLPVLMLNEGAFRCMYLLVMAIFWVTEALPLVVTSMIPVAPINGI 83
Db 11 KKLVLGLPLVAVPLLF--GPEYRCLFSIFLSTYTWIGEAFFIGVTSFLPLALYPILOI 68
QY 84 MSSDOTCRLYFKDITLVFMFGIMVALAVEYCNLHKRLALRVIQIVGCSPRRLHFLGIMVT 143
Db 69 VPSKQISPVYFKDSVLFVPLLF--GPEYRCLFSIFLSTYTWIGEAFFIGVTSFLPLALYPILOI 68
QY 144 MFLSMWISNACTAMMCPIIQAVLEEL--QAQGVCKINHE----POYQIVGNGKKNDEP 198
Db 119 --VSFFVSDTACTALMCPATAVALLMSMSDAVQHLKEDHKKPKPPDDATVAEKLIDDMT 176
QY 199 PYPTKITLCYLGIAVASLGGCGTIIIGTATNLTFKGIYEAFKNSSTQMDPPTFMFYSV 258
Db 177 PQDAGFCKALILACAHASLIGGTAITSTGPNLVFRENHKKYPGQVVTMTYLOMMVFAI 236
QY 259 PSMVLVYTLTLTFVFLQWHFMG----LW--RPKSKAEQVQRREGADVAKKVIDQR---- 307
Db 237 PMFVYLLASYIILVCYFMGPFSTFARWFERPESKEAH-----LKKLEIKNIQTM 285
QY 308 YKDLGEMSTHEIQVMILFTFMVVMYTRKPGIFLGWADLLNSKD--IRNSMPTIFVVMVCF 366
Db 286 YEDLDGVSNGEKS VVFFILLIGSMISRDGPTPGWGLLPHRNFTSDSVSGVLISCILF 345
QY 367 MLPANYAFURYCTRGGPV--PTGPTPSLITWKFIQTKVPWGLVFLVGGGFALAEKSGKSG 425
Db 346 VMPKD-----PPDPIIDPMAPILKWTDMKSKFSWSCTLLIGAGYAISEGVKSG 393
QY 426 MAKLG----NALIGLKVLPNSVLLVAVFLTAFSSNVAIANIIPVLAENSLAIE 481
Db 394 LSRLLSCGMKNIFVGNSSLP---LQLTVTIIVMTFASNVTSGSIFIPISLGAESMG 450
QY 482 IHPYLLPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITITLFFVFC 541
Db 451 VHPYLLPVTTCVACSFAMFLPISTPNNAVVDYTKVISVEMIVCGFLNACILITSLNM 510
QY 542 QTWGLVVPNLSFPE 557
|| : : : |||
```

```
Db 511 NTWTYFIF-SLNIPPE 525

RESULT 9
S30871
hypothetical protein R107.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S30871
R:Thomas, K.
submitted to the EMBL Data Library, July 1992
A:Reference number: S30871
A:Accession: S30871
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-539 <THO>
A:Cross-references: ENBL:Z14092
C:Genetics:
A:Genetic code: SGC4
A:Introns: 30/3; 80/3; 119/2; 217/2; 251/3; 293/2; 390/3; 416/1; 445/3; 491/3
A:Superfamily: sodium/sulfate cotransporter
C:Keywords: transmembrane protein

Query Match 24.4%; Score 730.5; DB 2; Length 539;
Best Local Similarity 30.9%; Pred. No. 1e-52;
Matches 170; Conservative 114; Mismatches 214; Indels 53; Gaps 12;

QY 24 KGLVVLPLLLCLPVLMLNEGAFRCMYLLVMAIFWVTEALPLVVTSMIPVAPINGI 83
Db 11 KKLVLGLPLVAVPLLF--GPEYRCLFSIFLSTYTWIGEAFFIGVTSFLPLALYPILOI 68
QY 84 MSSDOTCRLYFKDITLVFMFGIMVALAVEYCNLHKRLALRVIQIVGCSPRRLHFLGIMVT 143
Db 69 VPSKQISPVYFKDSVLFVPLLF--GPEYRCLFSIFLSTYTWIGEAFFIGVTSFLPLALYPILOI 68
QY 144 MFLSMWISNACTAMMCPIIQAVLEEL--QAQGVCKINHE----POYQIVGNGKKNDEP 198
Db 119 --VSFFVSDTACTALMCPATAVALLMSMSDAVQHLKEDHKKPKPPDDATVAEKLIDDMT 176
QY 199 PYPTKITLCYLGIAVASLGGCGTIIIGTATNLTFKGIYEAFKNSSTQMDPPTFMFYSV 258
Db 177 PQDAGFCKALILACAHASLIGGTAITSTGPNLVFRENHKKYPGQVVTMTYLOMMVFAI 236
QY 259 PSMVLVYTLTLTFVFLQWHFMG----LW--RPKSKAEQVQRREGADVAKKVIDQR---- 312
Db 237 PMFVYLLASYIILVCYFMGPFSTFARWFERPESKEAHKK-----LLEKNIQTMVEDLG 290
QY 313 PMSITHEIQVMILFTFMVVMYTRKPGIFLGWADLL--NSKDIRNSMPTIFVVMVCFMPLPAN 371
Db 291 DVSNGEKS VVFFILLIGSMISRDGPTPGWGLLPHSNFMSDSVSGVLISCILFVWPKD 350
QY 372 YAFURYCTRGGPV--PTGPTPSLITWKFIQTKVPWGLVFLVGGGFALAEKSGKSGMAKLI 430
Db 351 -----PPDPIIDPMAPILKWTDMKSKFSWSCTLLIGAGYAISEGVKSGLSLI 398
QY 431 G----NALIGLKVLPNSVLLVAVFLTAFSSNVAIANIIPVLAENSLAIEIHPY 486
Db 399 SCGMKNIFVGNSSLP---LQLTVTIIVMTFASNVTSGSIFIPISLGAESMGVHPY 455
QY 487 LILPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITITLFFVFCQTWGL 546
Db 456 LALPTTVACSFAMFLPISTPNNAVVDYTKVISVEMIVCGFLNACILITSLNNWTY 515
QY 547 VYENLSFPE 557
|| : : : |||
Db 516 FIF-SLNIPPE 525

RESULT 10
C89980
hypothetical protein SAI732 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
```

C;Accession: C89980  
R;Kuroda, M.; Ohca, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsubo, K.  
Lancet 357, 1225-1240, 2001  
A;Title: Whole genome sequencing of metacillin-resistant *Staphylococcus aureus*.  
A;Reference number: A89758; MUID:21311952; PMID:11418146  
A;Accession: C89980  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-520 <KUR>  
A;Cross-references: UNIPROT:Q99SX1; GB:BA000018; PID:g13701709; PIDN:BA043002.1; GSPDB:G  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: SAI732  
C;Superfamily: sodium/sulfate cotransporter

Query Match 22.6%; Score 676.5; DB 2; Length 520;  
Best Local Similarity 29.9%; Pred. No. 2.9e-48;  
Matches 164; Conservative 118; Mismatches 176; Indels 91; Gaps 14;

QY 18 FFANHWKLVVFLVPLCLPMLL--NEGAFRCMYLLV---MAIFWTEALPLVVTSM 72  
DB 28 FSAGQLICLI--LGPLELLTLLFFHPQDLFWKGVVLAITLWITWTEALPIAATSL 85

QY 73 IPIVAFPIGIMSSDOTCRLYFKDTLVFMGGMVALAVEYCNLHKLRLARVIOVGCSP 132  
DB 86 LPVILLPLGHLITPEQVSSEYNDIIFLGGFILAEMRWLHRTVALTIINLGAST 145

QY 133 RRLHFLGIMVTFMFLSMWISNAACATMPCPIIOAVLELOAQGVCKINHPQYQIVGNKK 192  
DB 146 SKILLGFVATGFLSMFVSNTAAVMIMPIGLAIKEA-----HDLQ-----E 188

QY 193 NNEDEPPYPTKITCYVILGIAYASLGGCGTIIIGTATNLTFKGIYEAFKNSTEQMDFFT 252  
DB 189 ANTNGTST-QKFEKSLVLAIGYAGTIGLGLTLPPLIILKGOYMQHFGH---EISFAK 244

QY 253 FMFYSVPSMLVYLLTTFVFLQW-----HFMLGRPKSKEAQVREGADVAKKVI 304  
DB 245 WMIVGIPVIVLLGITMILYLAFRHDLKYLPG-----GQTLI 283

QY 305 DQRYKDLGPMSEIHQVIMLFIEMVVMYFTRKPGIFLGWADLNSKDIRNSMP-----TIF 360  
DB 284 KQKLDLGMKYEKVVQTIPLVSLWITRE-----FLKKWEVTSVADGTIAIF 335

QY 361 VVMCFMPLPANYAFELVCTRGGVPPTPSLITWKFQIKVPMGLVFLGGFALAE 420  
DB 336 ISILLFIIPAK-----NTEKHRIIDWE-VAKELPWGVLLIFGGGLAKG 380

QY 421 SKQSGMAKLGNALIGLKVLPNSVLLVILVAVLTAFFSNVAIANIIPVLAEMSLAI 480  
DB 381 ISEGLAKWLGEQLKSLNGVSPILIVITVFLTEVTSNTATATMILPILATLSVAV 440

QY 481 EIHPLYLTPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIGPPIIITITLFPV 540  
DB 441 GVHPLLMAAANCAVMPLVGTTPNALIFGSGKISIKQMASVGFVWNLISAIIII-- 498

QY 541 CQTWGLVYV 549  
DB 499 -----LVVY 502

RESULT 11  
F64546  
sodium-dependent transporter - Helicobacter pylori (strain 26695)  
C;Species: Helicobacter pylori  
C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C;Accession: F64546  
R;Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997  
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A;Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A;Reference number: A64520; MUID:97394467; PMID:9252185  
A;Accession: F64546  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-552 <TOM>  
A;Cross-references: UNIPROT:O25003; GB:AE000541; GB:AE000511; NID:g2313299; PIDN:AA00728;  
C;Superfamily: sodium/sulfate cotransporter

Query Match 21.1%; Score 632; DB 2; Length 552;  
Best Local Similarity 31.9%; Pred. No. 1.5e-44;  
Matches 164; Conservative 94; Mismatches 186; Indels 70; Gaps 14;

QY 54 LVMAIFWTEALPLVVTSMIPVAFPIGIMSSDOTCRLYFKDTLVFMGGMVALAVEY 113  
DB 72 VLMGIWMTTEALDLPATALLPLVFSVSDQFASVSSYASPIIFLFGGFIILSMQK 131

QY 114 CNLHKLRLARVIOVGCSPRRLHFLGIMVTFMFLSMWISNAACATMPCPIIOAVLELOAQ 173  
DB 132 WNLHTRIALSIILLVGTSPRRLILGFMATGFLSMWVSNTATATVMMLPVGMVSL-QLVAK 190

QY 174 GYCKIN-----HEPO--YQIVGNKKN-----NEDEPPYPTKITCYVILGIAY 214  
DB 191 LVKEDASNSWQKEIEYKAHGIMSNIVHKDKITQVIOEKTIVRTNFSICLMGLIAY 250

QY 215 ASSLGGCGTIIIGTATNLTFKGIYEAFKNSTEQMDFFTFFMFYSVPSMLVYLLTTFVFLQW 274  
DB 251 AASIGSLGTLTTPNALLAGYMKTAFA---NIEIDFAQMMVFGTPLAFTMLILAWLLTY 307

QY 275 HFMLGRPKSKEAQVREGADVAKKVIDORYKDLGPMSEIHQVIMLFIEMVVMYFT 334  
DB 308 VIFPL-----KKEIPGKGE-----VIRVELKKGLRSLQAEISVGIIIFILASLGM-- 352

QY 335 RKPGIFLG-----WADLLNSKDIRNSMPTIFVVMCFMPLPANYAFELVCTRGGVPPTGP 389  
DB 353 -----IFLGMWLSWGVKIDKID---SVIANGVSALLFILPANHQGR----- 392

QY 390 TPLSLITWKFQIKVPMGLVFLGGFALAEKSGQMAKLGNALIGLKVLPNSVLLVY 449  
DB 393 ---LIDWG-VAKKLPMWVLLFGGGLALSAQFSKTLGSLWIGHLVSGFSLHPLFIIVMV 448

QY 450 ILVAVELTAFSSNVAIANIIPVLAEMSLAIEH-PLVLLPAGLACSMAFHLPVSTP 506  
DB 449 TLMVIELTETITNTATAAFLPVIGGVANGMVGHQSLLLTIPVALSATCAFMLPVTP 508

QY 507 PNALVAGYANIRTKDMAIAGIGPITIIITITLFPV 540  
DB 509 PNAIAGSGVVKITDMIKAGLMNLVGVVLISTF 542

## RESULT 12

E71961  
probable transporter - Helicobacter pylori (strain J99)

C;Species: Helicobacter pylori

A;Variety: strain J99

C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004

C;Accession: E71961

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*.

A;Reference number: A71800; MUID:99120557; PMID:9923682

A;Accession: E71961

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-548 <ARN>

A;Cross-references: UNIPROT:Q9ZMLB; GB:AE001458; GB:AE001439; NID:g4154713; PIDN:AA005784

A;Experimental source: strain J99

C;Genetics:

A;Gene: jhp0200

C;Superfamily: sodium/sulfate cotransporter

Query Match 20.9%; Score 626; DB 2; Length 548;

Qy	145	FLSMISNAAC	TAMCP	IIQAIVLBELOAQGVCKINHBEPQVQIVGNGKKNNEDEPPPYTKI	204
Dy		: : : :	: : : :	: : : :	:
Db	126	LISMISTNTSTTAMMLPLALGIL	-----AGV-	-----RETEREKVPFPVL	164
Qy	205	TLCYVIGTAGYASSLGCGCTTGATGNLTNLTFFGIYEARFKNSTEOMDPTTFEFYSVPMSLVY	264		
Dy		: : : :	: : : :	: : : :	:
Db	165	-----LGTAYSASVGIGITLVGSPPNGIAIGIL	-----GLSPFDMDKFGIPGVFLIL	210	
Qy	265	TLLTFVFLQWFMGLWRPEKSK-EAEQVQRGREGDAVAKVIDORYKDLGPMSIHETQVM	323		
Dy		: : : :	: : : :	: : : :	:
Db	211	FPLLFAILL-FL-VFRPSTDLKVERVQ-----EIKFEFTPQR-----VLVL	249		
Qy	324	LPIFWVMY-FTRKDCIPLGWADLLNSKDTPNSMPTIFVVVMCFMPLPANYAFURYCTRGG	382		
Dy		: : : :	: : : :	: : : :	:
Db	250	IFLTAALAWIFSFK-----IAPIFEVKKYFDTVVALLAVVALFI-	289		
Qy	383	GPVPTGPTSLITWKFIOTKPWGLVFELGGGFALAEBSQSGMAKLIGNALIG-LKVLP	441		
Dy		: : : :	: : : :	: : : :	:
Db	290	-----RLLDWRDVKEGVSWGTLLLFGGGIALSGIMKTGTAKFISOELVDVLHGLP	340		
Qy	442	NSVLLLVILVAVFATPSSNVAIANIIPVLAEMSLAIEHPLYLILPAGLCSMAFH	501		
Dy		: : : :	: : : :	: : : :	:
Db	341	TELELLITLVFVIFTELMSNTATTALIAPILPSTAQMICKPPEMLVIPAAVASCAFML	400		
Qy	502	PYSTPNALVAGYANIRTKDMAIGAGITTIITLFPV	540		
Dy		: : : :	: : : :	: : : :	:
Db	401	VPATPNAIVTGTGYIKQSQQMRVGLINIVFSIVLAAF	439		

RESULT 14  
164080  
probable membrane protein H10608 - Haemophilus influenzae (strain Rd KW20)  
C/Species: Haemophilus influenzae  
C/Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Jul-2004  
C/Accession: I64080  
R/Pfleschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Ke,  
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen,  
J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.;  
Science 269, 498-512, 1995  
A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.  
A/Title: Whole-genome random sequencing and assembly of Haemophilus influenzae  
A/Reference number: A64000; MUID:95350630; PMID:7542800  
A/Acession: I64080  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-461 <TIGR>  
C/Cross-references: UNIPROT:Q57486; GB:U32743; GB:I42023; NID:g1573597; PIDN:A  
C/Superfamily: probable transporter M30672

Query Match	14.9%	Score 447;	DB 2;	Length 461;
Best Local Similarity	24.2%;	Pred. 0.2.8e-29;		
Matches 127;	Conservative 115;	Mismatches 178;	Indels 104;	Gaps 16
QY	26	LVVFLVPLLCUPV-MLNLEGAEFFRCMYLLVMAIFWVTEALPLYVTS-MIPIVAFPMIGI	83	
Db	20	IVLFFVLLNVLPPFPKANSG-----LALLAFIAVLMLEALHTVITALLVPLIAV-ALGL	73	
QY	84	NSSDQTCRLYFKDITLVFMFGIMVALAVECNLHKRLALRVIOIVCCSPRRLHFGLIMVT	143	
Db	74	VSTRQALVGFADPTIFLFFGFGSLATLHTOKLQKLIANKMALARGNLFIAYVILFLIT	133	
QY	144	WFLSWISNAACTAMWCPIITQAVLEELQAOQVCKINHEPOYQIVGKNKKNDEPPYPTK	203	
Db	134	AFLSWMNNTATAMMLPLANGTILSQLDRE-----KQHTY-----	169	
QY	204	ITLCYVLGIAYASSLGGCGTIIGTATNLTPKGIYEARFKNSTEQMDPPTFMFVSPSMLV	263	
Db	170	--VFVLLGIIAYASISGGMGLTVGSPPAIVA-----SNLNLTFSDWLWGLPIMII	218	
QY	264	YTLITFFVLOHFNGLWRPKSKEAQEVQORREGADVAKKVIDORYKDLGPMIHEIQVMI	323	
Db	219	LLPLMIGILYIT-----IFPKP-----LHLNFQETPNI-----ENMPMRILIT	255	



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:39 ; Search time 76.9736 Seconds  
(without alignments)  
3805.323 Million cell updates/sec

Title: US-10-017-479A-2  
Perfect score: 2994  
Sequence: 1 MEIEIGEPQPPVKCSNPFA.....NSFPWAQIYAAALGNKTH 572  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt 03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2394	100.0	572	1	IND1 DROME	Q9vv22 drosophila
2	2394	100.0	590	2	Q7KUS6	Q7kus6 drosophila
3	1618.5	54.1	896	2	Q7QIT2	Q7qit2 anopheles g
4	1454.5	48.6	504	2	Q86B89	Q86b89 drosophila
5	1427	47.7	505	2	Q811B6	Q811b6 drosophila
6	1382.5	46.2	562	1	IND2 DROME	Q9vdd0 drosophila
7	1271	42.5	414	2	Q7PF67	Q7pf67 anopheles g
8	1042.5	34.8	587	2	Q35055	Q35055 rattus norv
9	1033.5	34.5	586	1	S132 MOUSE	Q9es88 mus musculu
10	1028	34.3	586	2	Q6AZR9	Q6azr9 xenopus lae
11	1026	34.3	592	1	S132 HUMAN	Q11183 homo sapien
12	1024.5	34.2	624	2	Q64ZNS	Q64zn5 xenopus tro
13	1024	34.2	593	1	S132 RABIT	Q28615 oryctolagus
14	1023.5	34.2	613	2	Q803K7	Q803k7 brachydanio
15	1022.5	34.2	613	2	Q6P3L2	Q6p3l2 brachydanio
16	1010.5	33.8	622	2	Q57661	Q57661 xenopus lae
17	1007	33.6	605	2	Q866R0	Q866r0 didelphis m
18	1004	33.5	568	2	Q86YT5	Q86yt5 homo sapien
19	1002	33.5	568	2	Q6ZMGI	Q6zmg1 homo sapien
20	993.5	33.2	587	1	S132 RAT	P70545 rattus norv
21	983	32.8	572	2	Q8CJ44	Q8cj44 rattus norv
22	975.5	32.6	600	1	S133 RAT	Q9z025 rattus norv
23	972.5	32.5	600	1	S133 MOUSE	Q9ly63 mus musculu
24	965.5	32.2	602	1	S133 HUMAN	Q8wvt9 homo sapien
25	964.5	32.2	602	2	Q8IVB1	Q8ivb1 homo sapien
26	948.5	31.7	572	2	Q67BT3	Q67bt3 mus musculu
27	910	30.4	596	2	Q6NWY4	Q6nwy4 brachydanio
28	901.5	30.1	601	2	Q7ZWL2	Q7zwl2 xenopus lae
29	893	29.8	602	2	Q6DJN7	Q6djn7 xenopus lae
30	892.5	29.8	594	1	S131 MOUSE	Q9jhi4 mus musculu
31	890.5	29.7	594	2	Q80YB5	Q80yb5 mus musculu

32	889.5	29.7	627	1	S134 HUMAN	Q9uk94 homo sapien
33	888	29.7	602	2	Q6E7G8	Q6e7g8 xenopus lae
34	885	29.6	595	1	S131 RAT	Q07782 rattus norv
35	885	29.6	626	2	Q8N631	Q8n631 homo sapien
36	883.5	29.5	625	2	Q8BZ82	Q8bz82 mus musculu
37	879.5	29.4	595	1	S131 HUMAN	Q9bzw2 homo sapien
38	878.5	29.3	583	2	Q6NY28	Q6ny28 brachydanio
39	878.5	29.3	583	2	Q6PE27	Q6pe27 pseudopleur
40	876	29.2	601	2	Q6W7I2	Q9w7i2 caenorhabdi
41	874.5	29.2	582	1	NAD1 CAEEL	Q93655 caenorhabdi
42	838.5	28.0	619	2	Q6L970	Q6l970 anguilla ja
43	819	27.4	520	2	Q8N8K4	Q8n8k4 homo sapien
44	812	27.1	577	1	YK66 CAEEL	P46556 caenorhabdi
45	799.5	26.4	551	1	NAD3 CAEEL	Q21339 caenorhabdi
46	791.5	26.4	551	1	NAD2 CAEEL	P32739 caenorhabdi
47	695	23.2	552	2	Q8NEE3	Q8ene3 oceanobacil
48	687	22.9	546	2	Q65NCO	Q65nc0 bacillus li
49	676.5	22.6	520	2	Q8NV55	Q8nv55 staphylococ
50	676.5	22.6	520	2	Q99SX1	Q99sx1 staphylococ
51	676.5	22.6	520	2	Q7A4P8	Q7a4p8 staphylococ
52	676.5	22.6	520	2	Q6G816	Q6g816 staphylococ
53	673.5	22.5	520	2	Q6GFE0	Q6gfe0 staphylococ
54	644	21.5	626	2	Q6AQ77	Q6aq77 desulfotale
55	632	21.1	552	2	Q25003	Q25003 helicobacte
56	626	20.9	548	2	Q9ZML8	Q9zml8 helicobacte
57	621.5	20.8	540	2	Q8LG88	Q8lgh8 arabidopsis
58	620.5	20.7	450	2	Q75MH3	Q75mh3 homo sapien
59	619.5	20.7	540	2	Q9MAW4	Q9maw4 arabidopsis
60	611	20.4	510	2	Q7UUK9	Q7uuk9 rhodospirell
61	610.5	20.4	510	2	Q8NTS7	Q8nts7 corynebacte
62	610.5	20.4	527	2	Q6M8D0	Q6m8d0 corynebacte
63	609.5	20.4	524	2	Q6NK59	Q6nk59 corynebacte
64	607.5	20.3	495	2	Q67T81	Q67t81 symbiobacte
65	607.5	20.3	540	2	Q9FEH9	Q9feh9 arabidopsis
66	602	20.1	462	2	Q9FGK1	Q9fgk1 arabidopsis
67	601.5	20.1	474	2	Q65MM0	Q65mm0 bacillus li
68	596	19.9	587	2	Q8FU27	Q8fu27 corynebacte
69	586.5	19.6	548	2	Q6ZFH7	Q6zfh7 oryza sativ
70	580	19.4	450	2	Q8FMZ5	Q8fmz5 corynebacte
71	528.5	17.7	302	2	Q68D44	Q68d44 homo sapien
72	511	17.1	446	2	Q66449	Q66449 aquifex aeo
73	485.5	15.5	463	2	Q9CLQ0	Q9clq0 pasteurella
74	459	15.3	459	2	Q8T3Y2	Q8t3y2 drosophila
75	450.5	15.0	470	2	Q65R88	Q65r88 manheimia
76	447	14.9	461	1	Y608 HAEIN	Q57486 haemophilus
77	446	14.9	456	2	Q6LQW4	Q6lqm4 photobacter
78	445.5	14.9	471	2	Q9JV43	Q9jv43 neisseria m
79	444.5	14.8	471	2	Q9K032	Q9k032 methanobact
80	440.5	14.7	443	2	Q26881	Q26881 neisseria m
81	431.5	14.4	461	2	Q7N822	Q7n822 photorhabdu
82	430.5	14.4	456	2	Q87Q94	Q87q94 vibrio para
83	430.5	14.4	474	2	Q8GF70	Q8gf70 photorhabdu
84	426.5	14.2	462	2	Q9KNE0	Q9kne0 vibrio chol
85	424	14.2	459	2	Q66615	Q66615 yersinia pe
86	420.5	14.0	456	2	Q7MLG6	Q7mlg6 vibrio vuln
87	420.5	14.0	456	2	Q8D908	Q8d908 vibrio vuln
88	420	13.9	451	2	Q7M8Y1	Q7m8y1 wolinnella e
89	417	13.9	459	2	Q8CZT3	Q8czt3 yersinia pe
90	411	13.7	456	2	Q8ZHX0	Q8zhx0 yersinia pe
91	411	13.7	463	2	Q65SS5	Q65ss5 manheimia
92	409.5	13.7	464	2	Q9AED9	Q9aed9 leptospira
93	409.5	13.7	464	2	Q72QJ5	Q72qj5 leptospira
94	409	13.7	464	2	Q8VTX1	Q8vtx1 leptospira
95	405.5	13.5	464	2	Q8F5L4	Q8f5l4 leptospira
96	403.5	13.5	464	2	Q9S4F3	Q9s4f3 leptospira
97	400	13.4	464	2	Q9ZGJ3	Q9zgj3 leptospira
98	391.5	13.1	432	1	Y672 METJA	Q58085 methanococc
99	390.5	13.0	471	2	Q6LSU9	Q6lsu9 photobacter
100	383	12.8	487	2	Q9KSE3	Q9kse3 vibrio chol

ALIGNMENTS

RESULT 1  
ID IND1 DROME STANDARD; PRT; 572 AA.  
AC Q9VVT2; Q9NHY9;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 25-JAN-2005 (Rel. 46, Last annotation update)  
DE I'm not dead yet protein (IND1 transporter protein) (drIndy).  
GN Names=Indy; ORFNames=CG3979;  
OS Drosophila melanogaster (fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophiliidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RN SEQUENCE FROM N.A., AND FUNCTION.  
RP MEDLINE=22250702; PubMed=12186628; DOI=10.1042/BJ20021132;  
RA Inoue K., Fei Y.-J., Huang W., Zhuang L., Chen Z., Ganapathy V.;  
RT "Functional identity of Drosophila melanogaster Indy as a cation-  
RT independent, electroneutral transporter for tricarboxylic acid-cycle  
RT intermediates";  
RL Biochem. J. 367:313-319 (2002).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Berkely;  
RX MEDLINE=2019606; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delecher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleby J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195 (2000).  
RN [3]  
RN GENOME REANNOTATION.  
RP MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,

RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RT systematic review";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).  
RN [4]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Berkely; TISSUE=Embryo;  
RX MEDLINE=22426066; PubMed=12537569;  
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
RA George R.A., Guarin H., Krommiller B., Pacleby J.M., Park S., Wan K.H.,  
RA Rubin G.M., Celniker S.E.;  
RT "A Drosophila full-length cDNA resource";  
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8 (2002).  
RN [5]  
RN SEQUENCE OF 366-572 FROM N.A.  
RP STRAIN=Canton-S;  
RX MEDLINE=20050669; PubMed=10581279;  
RA Schmid K.J., Nigro L., Aquadro C.F., Tautz D.;  
RT "Large number of replacement polymorphisms in rapidly evolving genes  
RT of Drosophila. Implications for genome-wide surveys of DNA  
RT polymorphism";  
RL Genetics 153:1717-1729 (1999).  
RN [6]  
RN IDENTIFICATION, FUNCTION, AND TISSUE SPECIFICITY.  
RP PubMed=1118146; DOI=10.1126/science.290.5499.2137;  
RA Rogina B., Reenan R.A., Nilsen S.P., Helfand S.L.;  
RT "Extended life-span conferred by cotransporter gene mutations in  
RT Drosophila";  
RL Science 290:2137-2140 (2000).  
RN [7]  
RN FUNCTION, SUBCELLULAR LOCATION, INDUCTION, AND TISSUE SPECIFICITY.  
RP PubMed=12391301; DOI=10.1073/pnas.222531899;  
RA Knauf F., Rogina B., Jiang Z., Aronson P.S., Helfand S.L.;  
RT "Functional characterization and immunolocalization of the transporter  
RT encoded by the life-extending gene Indy";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:14315-14319 (2002).  
RN [8]  
RN FUNCTION: Cation-independent electroneutral transporter (not  
RN associated with membrane depolarization) of a variety of  
RN tricarboxylic and dicarboxylic acid-cycle intermediates. There is  
RN also small, but detectable, transport of monocarboxylics.  
RN Transport is through the epithelium of the gut and across the  
RN plasma membranes of organs involved in intermediary metabolism and  
RN storage. Affinity for substrates is citrate > succinate >  
RN pyruvate. Fumarate, a-ketoglutarate, and glutamate are also  
RN transported, but not lactate. Transport mechanism that is not  
RN coupled to Na(+), K(+), or Cl(-). Function is shown in Xenopus  
RN oocytes and human retinal pigment epithelial (HRPE) cell lines.  
RN -I- SUBCELLULAR LOCATION: Integral membrane protein; basolateral  
RN membrane of cells in the midgut.  
RN -I- TISSUE SPECIFICITY: In adults, abundantly expressed in the fat  
RN body, basolateral region of midgut cells and oenocytes. Low level  
RN expression is seen in the halteres, procardia, restricted regions  
RN of the esophagus and hindgut, base of the legs and in a subset of  
RN cells in the third segment of the antennae.  
RN -I- INDUCTION: Completely inhibited by DIDS. Modest but significant  
RN inhibition by phloretin or furosemide.  
RN -I- MISCELLANEOUS: The life-extending effect of mutations is likely  
RN caused by an alteration in energy balance caused by a decrease in  
RN transport function.  
RN -I- SIMILARITY: Belongs to the SLC13A transporter family. NACD  
RN subfamily.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----

EMBL; AF509505; AAN86815.1; -;  
EMBL; AE003519; AAP49226.1; -;



RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
a genomic perspective.";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Milera S., Croesby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
systematic review.";  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
RN EMBL: AE003519; AAF49227.2; -.  
DR GO: GO:0016020; C:membrane; IEA.  
DR GO: GO:0005215; F:transporter activity; IEA.  
DR GO: GO:0006814; P:sodium ion transport; IEA.  
DR InterPro: IPR001898; Na/sul. symport.  
DR Pfam: PF00939; Na sulph. sympt. 1.  
DR PROSITE: PS01271; NA SULFATE; UNKNOWN 1.  
SQ SEQUENCE 590 AA; 65585 MW; ECA7857BB6B34414 CRC64;  
  
Query Match 100.0%; Score 2994; DB 2; Length 590;  
Best Local Similarity 100.0%; Pred. No. 2.8e-216;  
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MEIEIGQOPPPVKCSNFFANHWKGLVFLVLLCLPVLNLSGAEFRMYLLVNAIFW 60  
DB 19 MEIEIGQOPPPVKCSNFFANHWKGLVFLVLLCLPVLNLSGAEFRMYLLVNAIFW 78  
  
QY 61 VTEALPLYVTSMPIVAFPIGIMSSDQTCRLYFKDITLVNFMGIMVALAVEYCNLHKRL 120  
DB 79 VTEALPLYVTSMPIVAFPIGIMSSDQTCRLYFKDITLVNFMGIMVALAVEYCNLHKRL 138  
  
QY 121 ALRVIQVGSRRLLHGLIMVTWFLSMWISNAACTAMMCPITQAVLEELQAGVCKINH 180  
DB 139 ALRVIQVGSRRLLHGLIMVTWFLSMWISNAACTAMMCPITQAVLEELQAGVCKINH 198  
  
QY 181 EPQYQIVGGNKKNEDEPPYPTKITLCYLGIAVASSLGCGTIGTATNLTFGKIYEAR 240  
DB 199 EPQYQIVGGNKKNEDEPPYPTKITLCYLGIAVASSLGCGTIGTATNLTFGKIYEAR 258  
  
QY 241 FKNSTEQMDPTTFMFYSPVSMVYLLTTFVLOWHFMGLWRPKSKEAQVQREGADV 300  
DB 259 FKNSTEQMDPTTFMFYSPVSMVYLLTTFVLOWHFMGLWRPKSKEAQVQREGADV 318  
  
QY 301 KKVIDQRYKDLGPNMSIHEIQMILFIPMVVMTFRKPGIFLGWADLLNSKDINSMTIF 360  
DB 319 KKVIDQRYKDLGPNMSIHEIQMILFIPMVVMTFRKPGIFLGWADLLNSKDINSMTIF 378  
  
QY 361 VVWCMFLPANYAFRLYCTRGGVPVPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420  
DB 379 VVWCMFLPANYAFRLYCTRGGVPVPTPSLITWKFIQTKVPWGLVFLGGGFALAE 438  
  
QY 421 SKOSGMKLIQNALIGLKVLPNSVLLVILVAVFLTAFSSNVAIANIIPVLAESLAI 480  
DB 439 SKOSGMKLIQNALIGLKVLPNSVLLVILVAVFLTAFSSNVAIANIIPVLAESLAI 498  
  
QY 481 EIHPLYLILPAGLACSMAFHPLVSTPPNALVAGYANIRTKDMAIGPTIIITILFVF 540  
DB 499 EIHPLYLILPAGLACSMAFHPLVSTPPNALVAGYANIRTKDMAIGPTIIITILFVF 558

QY 541 CQTWGLVVPNLNSPPEWAQIYVAAAALGNKTH 572  
DB 559 CQTWGLVVPNLNSPPEWAQIYVAAAALGNKTH 590  
  
RESULT 3  
QYQIT2 PRELIMINARY; PRT; 896 AA.  
ID Q7QIT2  
AC Q7QIT2; 01-MAR-2004 (TEMBLrel. 26, Created)  
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)  
DE EbiP7623 (Fragment).  
GN Name=ebiG7623; ORFNames=ENSANGG00000005753;  
OS Anopheles gambiae str. PEST.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.  
OX NCBI\_TaxID=180454;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PEST;  
RA Anopheles Genome Sequencing Consortium;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
-!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL: AAB0100807; EAA03999.1; -.  
DR GO: GO:0016020; C:membrane; IEA.  
DR GO: GO:0005215; F:transporter activity; IEA.  
DR GO: GO:0006814; P:sodium ion transport; IEA.  
DR InterPro: IPR001898; Na/sul. symport.  
DR Pfam: PF00939; Na sulph. sympt. 4.  
DR PROSITE: PS01271; NA SULFATE; UNKNOWN 1.  
FT NON\_TER 1 1  
FT NON\_TER 896 896  
SQ SEQUENCE 896 AA; 98460 MW; D59FAE475FEE7982 CRC64;  
  
Query Match 54.1%; Score 1618.5; DB 2; Length 896;  
Best Local Similarity 55.6%; Pred. No. 7.1e-113;  
Matches 298; Conservative 102; Mismatches 111; Indels 25; Gaps 3;  
  
QY 24 KGIWVFLVLLCLPVLNLSGAEFRMYLLVNAIFWVTEALPLYVTSMPIVAFPIGMI 83  
DB 386 QSLAVLLVLLIAASVFLVDTSPPCWYIWMVGVFTEALPLITSMPLVFLPLMGI 445  
  
QY 84 MSDDQTCRLYFKDITLVNFMGIMVALAVEYCNLHKRLALRVIQVGSRRLLHGLIMVT 143  
DB 446 LDTNRTCMYMKDITLVNFMGIMVALAVEYCNLHKRLALRVIQVGSRRLLHGLIMVT 505  
  
QY 144 MFLSMWISNAACTAMMCPITQAVLEELQAGVCKINHPEQYQIVGGNKKNEDEPPYPTK 203  
DB 506 MFLSMWISNAACTAMMCPITQAVLEELQAGVCKINHPEQYQIVGGNKKNEDEPPYPTK 563  
  
QY 204 ITLCYLGIAVASSLGCGTIGTATNLTFGKIYEARFKNSTEQMDPTTFMFYSPVSMV 263  
DB 564 ITWCYIIGTAYAAATCGVGTIVSGVNLTKGIYESRFPQA-PGIDPPSPMFINPGMLL 622  
  
QY 264 YTLTTFVLOWHFMGLWRPKSKEAQVQREGADVAKKVIDORYKDLGPNMSIHEIQMIL 323  
DB 623 FTFLTWMLQWLFMGLFRPNSEDAARAANTGPEGEAVERVIENRLELGMTSHEIGVAF 682  
  
QY 324 LFIFMVVMTFRKPGIFLGWADLLNSKDINSMTIFVWVCMFLPANYAFRLYCTRGG 383  
DB 683 FFLSVLVTFTREGFMGMWADLVDPVKYKDATPALFVIMLFIIVPADWNLCLAFFHKG 742  
  
QY 384 PVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAEKSGKSGMAKLIQNALIGLKVLPNS 443  
DB 743 -----NLLFLGGGFALAEKSGRTSGMSALIGQSLAGLKVLPPL 780  
  
QY 444 VLLVAVFLTAFSSNVAIANIIPVLAESLAIETHPLYLILPAGLACSMAFHPLV 503  
DB 781 LLLFLVGVGEFTEFTSTNVAICNVVLPVLAETIAIEIHPFLYMPVAMVCSFSLPV 840

QY 504 STPPNALVAGYANIRTKDMAIGAGTITITITLTFVFCQWGLVYVYNLNSPPWA 559  
 Db 841 GTPPNAIVAGVGNIRTKDMAVAGIGPSIFTLLISWASFTWGSIIYPELASFPDWA 896

RESULT 4  
 Q86B89 PRELIMINARY; PRT; 504 AA.  
 AC Q86B89;  
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE CG32921-PB (CG32921-pb)  
 GN ORFNames=CG32921;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaesner K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lascko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang X., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milchena N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195 (2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
 RA Pacleeb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
 melanogaster euchromatic genome sequence."  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).

RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 a genomics perspective."  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Battencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 systematic review."  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003728; AAC41577.1; -  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005215; P:transporter activity; IEA.  
 DR GO; GO:0006814; P:sodium ion transport; IEA.  
 DR InterPro; IPR001898; Na/sul\_sympot.  
 DR Pfam; PF00939; Na\_sulph\_symp; 2.  
 SQ SEQUENCE 504 AA; 5386 MW; 28FDBFCE65065F4A CRC64;  
 Query Match 48.6%; Score 1454.5; DB 2; Length 504;  
 Best Local Similarity 52.4%; Pred. No. 8e-101;  
 Matches 268; Conservative 99; Mismatches 135; Indels 9; Gaps 4;  
 QY 50 MYLLVMAIFWVTEALPLYVTSMIPVAFPPFMGIMSSDQTCRLYFKDTLVNFMGIMVAL 109  
 Db 1 MYLVANMALFWITEALPLYLTALFVWVFLPLFGLTSEKVSFYSFSDVVMFVIGLLIAL 60  
 QY 110 AVECNLHKLRLALVIQVCSPPRLHGLIWMFLSMWISNACTAMCPIQAVLEB 169  
 Db 61 AIEYSLNHLQIALNTILVGCSPRLHGLVWVTCFISLWISNSAANTAMCPIKAVLNE 120  
 QY 170 LOAGGVCKINHEPOYQIVGGKKNEDPPVPTKITLCYVLGIAYVASSIGCGTIGTAT 229  
 Db 121 METQNIFAIYKTQEEPV-----EGDPPHPTSTISMAFYGIAYSSIGCGTIGTAT 174  
 QY 230 NLTFKGIYEARFNSTQMDPPTFMFVSVP-SMLVYTLTLLTFVFLQWHFMGLWRPKSKEAQ 288  
 Db 175 NLTYGLYDTRFRNSDEKIDFPIFMAYSVVPVVLVLFVLFVSLQVTHMGLFRNSKIGQ 234  
 QY 289 EVQGRREGADVAKVVDQRYKIDGPMISHEIQVLMILFVWVVFTRKPGIFLGWADLLN 348  
 Db 235 EVKGAESQDVVKDVIKQKAEALGPMSCHEIQVGLLFVLMIFLLFTRKPGFPGWADFLN 294  
 QY 349 SKDIRNSMPTIFVVMCMFLPANYAFRLYCTRRGGVPVPTGPTSLITWKEITQKVPWGLV 408  
 Db 295 ANAGSGPPVFFATILLPALPTQTFPKYCGK-APFPGQTLDCLSVYCYCKYTPFGLA 353  
 QY 409 FLLGGGFALAEKSGQSGMAKILGNALIGLKLPLNSVLLVLLVAVFLTAFFSSNVAIANI 468  
 Db 354 FLL-GGFALAEKSGVSGMAKLGESLAFAGEMHSLVISMCIILSLCTAFASNAICNI 412  
 QY 469 IIPVLAEMSLAIEIHPYLYLILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIGAG 528  
 Db 413 LIPIFSEMALAIEVHPMKLTFFPALACSLAFHLFPVSTPPNALISGFTGLTKYMAIAGIL 472

```
QY 529 PTIITITLTFVFCQWGLVVPNLSPPEWA 559
Db 473 PTCWAFCLLFTGTGVTMTLTPGTTFPSWA 503

RESULT 5
Q811B6 PRELIMINARY; PRT; 505 AA.
ID Q811B6
AC Q811B6;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE IndY-2-PA.
GN Name=Indy-2;
OS Drosophila pseudoobscura (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7237;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tucson 14011-0121.4;
RA Bergman C.M., Pfeiffer B.D., Rincon-Limas D.E., Hoskins R.A., Park S.,
RA Gairke A., Mungall C.J., Wang A.M., Kronmiller B., Pacleb J., Park S.,
RA Stapleton M., Wan K., George R.A., de Jong P.J., Botas J., Rubin G.M.,
RA Celniker S.E.;
RT "Assessing the impact of comparative genomic sequence data on the
RT functional annotation of the Drosophila genome."
RL Genome Biol. 3:research0086-research0086(2002).
DR EMBL: AY190949; AAC01056.1; -
DR FlyBase: FBgn0064420; Dpse\Indy-2.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006814; F:sodium ion transport; IEA.
DR InterPro: IPR001395; Aldo/ket red.
DR Dfam: PF00939; Na sulph symd; 2.
DR PROSITE: PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN 1.
SQ SEQUENCE 505 AA; 55608 MW; C1F8E3D37C579A40_CRC64;

Query Match 47.7%; Score 1427; DB 2; Length 505;
Best Local Similarity 52.4%; Pred. No. 9.3e-99;
Matches 271; Conservative

QY 50 MYLLVMAIFWVTEALPLYTSMIPVAFPIMGIMSSDQCRLYFKDLYVFMGGINVAL 109
Db 1 MYLVSNMALFWITEAIPLYLTSLFPVVFLEPLFDILGSDQVCKLYFSVTVMFGLLIAL 60
QY 110 AVEYCNLHKRLARVIOVGCSPRLHFLGLIMVTMFLSMWISNAACATMMCPITIOAVLEE 169
Db 61 AIEYVSNLHQRIAMKTIILVGCSPRLHFLGLVWVTCFISLWISNATAMCPIVKAVLSE 120
QY 170 LOAGGVCKI-----NHEPQYQIVGKNKKNDEPPYPTKITLCYVLGIYASSLGGCGTII 225
Db 121 LDSQNIITVYKSEEPW-----EGDPPHPSTISMAFYGVAYAAATIGGCGTII 170
QY 226 GTATNLTFKGIYARFNKNSTQMDPFTFMFVSFVSMVLYLLTFV-LQHFVGLWRPKS 284
Db 171 GTGTNLTFKGLYDRFRPKSKTQIDRIFEMAYAIPIVVIVNVLVLLYFSLQTHMGLFRGKS 230
QY 285 KEAEVORGREGADVAKVVDORVKDLPMSIHBIQWMLFIFVMVVFTRKPGIFLGWA 344
Db 231 QTGLEVRGTGEGQAVTVTVIKARHQELQPMTCHEIQVTVILFVLLFTRKPGFVTGWG 290
QY 345 DLLNSKDIRNSMPTIFVVMCMFLPANYAFURYCTRGGVPTGPT-PSLITWKFIQTKV 403
Db 291 DFLNAQIGSGPPVWLPVLLFALPTQYTFKYCC--GKPPFTQSQMDALLSWIYIHRNT 348
QY 404 PWGLVFLGGCFALAEKSGMAKLGNALIGLKVLPNSVLLLVTLVAVELTAFFSSNV 463
Db 349 PWGLCFLL-GGFALAEKSGMAKLGLSLKFAASKMPPIVVEGMCILIGLICTAFSSNV 407
```

```
QY 464 ATANIIIPVLAEMSLAIEHPLYLILPAGLAGSMAPHLFVSTPPNALVAGYANIRTKDMA 523
Db 408 AICNLILPIFSEMAIAIKHPLKLTLPSPSLASMAVHLFVSTPPNAIISYAGIKTKYLA 467
QY 524 IAGIGPTIITITLTFVFCQWGLVVPNLSPPEWAQ 560
Db 468 LAGILPTIAFWIVLMLNSQ-YGNIIFPESTKFPDWAQ 503

RESULT 6
IND2_DROME
ID IND2_DROME STANDARD; PRT; 562 AA.
AC Q9VD00; Q95T83;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE I'm not dead yet protein 2.
GN Name=Indy-2;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2195-2195(2000).
RN [2]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
```

Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.;  
"Annotation of the Drosophila melanogaster euchromatic genome: a systematic review";  
Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
[3]  
SEQUENCE FROM N.A.  
STRAIN=Berkeley; TISSUE=Head, and Testis;  
MEDLINE=2242606; PubMed=12537569;  
Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
Garage R.A., Guarin H., Konniger B., Pacleb J.M., Park S., Wan K.H.,  
Rubin G.M., Celniker S.E.;  
"A Drosophila full-length cDNA resource";  
Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
[4]  
IDENTIFICATION.  
PubMed=1118146; DOI=10.1126/science.290.5499.2137;  
Rogina B., Reenan R.A., Nilsen S.P., Helfand S.L.;  
"Extended life-span conferred by cotransporter gene mutations in Drosophila";  
Science 290:2137-2140(2000).  
-I- FUNCTION: Cation-independent electroneutral transporter (not associated with membrane depolarization) of a variety of tricarboxylic and dicarboxylic acid-cycle intermediates. There is also small, but detectable, transport of monocarboxylates. Transport is through the epithelium of the gut and across the plasma membranes of organs involved in intermediary metabolism and storage (By similarity).  
-II- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
-I- SIMILARITY: Belongs to the SLC13A transporter family. NADC subfamily.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AE003728; AAN13819.1; -  
ENBL; AV060287; AL25326.1; -  
EMBL; BT001329; AAN71084.1; -  
FlyBase; FBgn0011603; Indy-2.  
GO; GO:0005887; C:integral to plasma membrane; ISS.  
GO; GO:0015137; C:intracellular transporter activity; ISS.  
GO; GO:0050933; P:pyruvate transporter activity; ISS.  
GO; GO:0015411; P:succinate transporter activity; ISS.  
GO; GO:0015746; P:citrate transport; ISS.  
GO; GO:0008340; P:determination of adult life span; ISS.  
GO; GO:0006948; P:pyruvate transport; ISS.  
GO; GO:0015744; P:succinate transport; ISS.  
InterPro; IPR001898; Na/sul\_symport.  
Pfam; PF00939; Na\_sulph\_sympt; 1.  
PROSITE; PS01271; NA\_SULFATE; FALSE\_NEG.  
KW Transmembrane; Transport.  
FT TRANSMEM 24 44 Potential.  
FT TRANSMEM 64 84 Potential.  
FT TRANSMEM 93 113 Potential.  
FT TRANSMEM 135 155 Potential.  
FT TRANSMEM 200 220 Potential.  
FT TRANSMEM 249 269 Potential.  
FT TRANSMEM 308 328 Potential.  
FT TRANSMEM 350 370 Potential.  
FT TRANSMEM 393 413 Potential.  
FT TRANSMEM 430 450 Potential.  
FT TRANSMEM 478 498 Potential.  
FT TRANSMEM 515 535 Potential.  
SQ SEQUENCE 562 AA; 61934 MW; D9F04C05FAA73ED3 CRC64;



Qy 478 LAIEHPLYLILPAGLACSMAPHLIVSTPPNALVAGYANIRTKDMAIAGIGFTIITL 537  
 Db 487 QAICLHPYVNLPCCTLAASLAFMLFVATPPNAIVFSGGLKYSVMARAGFLNLIIGVLAI 546  
 Qy 538 VFVQCTGVLVVVNLNSPFEWAQIVAAAAAGNKTH 572  
 Db 547 TLSINSWSIPIF-KLDTFPSWAHSNTSQCLLNPSN 580

RESULT 9  
 S132 MOUSE STANDARD; PRT; 586 AA.  
 AC Q95S88;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Solute carrier family 13, member 2 (Renal sodium/dicarboxylate cotransporter) (Na(+)/dicarboxylate cotransporter 1) (NaDC-1).  
 GN Name=SLC13a2; Synonyms=Nadcl, Sdcl1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=20425391; PubMed=10966927;  
 RA Pajor A.M., Sun N.N.;  
 RT "Molecular cloning, chromosomal organization and functional characterization of a sodium/dicarboxylate cotransporter from mouse kidney.";  
 RT Am. J. Physiol. 279:F482-F490(2000).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., McQuellan N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettmann M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Cotransport of sodium ions and dicarboxylates such as succinate and citrate.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- TISSUE SPECIFICITY: Highly expressed in kidney and small intestine. Not detectable in brain, heart, stomach and skeletal muscle.  
 CC -1- SIMILARITY: Belongs to the SLC13A transporter family. NADC subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announcement/) or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch).

[illegible]

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
ID S132a2-prov protein.  
GN Name=s132a2-prov;  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RL initiative.";  
RL Dev. Dyn. 225:384-391 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshitoki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
RA Klein S., Gerhard D.S.;  
RL EMBL; BC077435; AAH77435.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006814; P:sodium ion transport; IEA.  
DR InterPro; IPR001898; Na/sul\_sympot.  
DR Pfam; PF00939; Na\_sulph\_symp; 1.  
DR PROSITE; PS01271; NA\_SULFATE; 1.  
SQ SEQUENCE 586 AA; 64959 MW; 2E3646A1D81F31DB CRC64;

Query Match 34.38; Score 1028; DB 2; Length 586;  
Best Local Similarity 37.18; Pred. No. 9.5e-69;  
Matches 221; Conservative 121; Mismatches 187; Indels 66; Gaps 11;

Qy 15 CSNFFANHKGVLVFLVPLCLPVLNMEGAEFRCMYLLLVMAIFVTEALPLVVTSMIP 74  
Db 4 CMKWLNLCKRFLILFWTPLLPLLLPLLIQTKEASCAVVIIMAVYWCTEVPLAVTALLP 63

Qy 75 IVAPFMINGMSSDQTCRLYKFDLTPMFGIMVALAVEYNLHKRLALRVIQVGCSPRR 134  
Db 64 VLLFPFMGILESCKVCQYKLDNTNMFVGLIVAVAEQWNLHKRIALKVLLIVGVRLPAL 123

Qy 135 LHFGILVWTFSLWISNACTAMWCPIIOAVLEEL----- 170  
Db 124 LMLGFMGVTAFLSWISNTATTAMVFIQAVLSQLTASEDPSMLESQNTNPAALELE 183

Qy 171 -----QAQGVCKINHEPQYQVGGNKNQNEDEPPYPTK-ITLCYYLGIAYASSLGGCG 222

Db 184 TKNAIPMLQVQTVSNHGVPEVP-----EEVPEKDERKHTSKGMLC-----VCYAAISGGTA 235  
Qy 223 TIIGTATNLTFKGIYEAFKNSQTEQMDFTTFMFYSVPSMLVYLLTFVFLQWFMGLWRP 282  
Db 236 TLTGTGPNVLVKGQFTQIFPNNGDILFNASWFGAFPNVMVIMCLSWFWLQFSFGFNFK 295  
Qy 283 KS-----KEAQEVQRGREGADVAKKVIDQRYKOLGPMHSIHIEIQVMILFIFVNVVYFRKPG 338  
Db 296 KTWGCGATASEKER-----AAYSVIREYRKLGPISYAESSVLFIILLVLLWFTDFPG 349  
Qy 339 IFLGWADLLNSKD-----INSMPTIFVVMVCMFLPANYAFRLRYCVRGG-----PV 385  
Db 350 FVVGWATILFNKDNIEYATDATVAVFAVFLFLLPATKPKFGCKKARNSFDLEDFEEQK 409  
Qy 386 PTGPTPSLTITWKFIQKVPWGLVFLGGGFGALAGSKQSGMAKLIGNALIGLKVLPN-SV 444  
Db 410 DTFFSAPLLTWKVVQKQMPWSIVLLGGGFGALAGSDAGSLSHWLGQQTPLHSIPWAI 469  
Qy 445 LLLVVLVAVFLTAFSSNVAIANIIPVLAEMSLAIEIHPYLILPAGLACSMFAHLPVS 504  
Db 470 AILSLMIAVF-TECASNVATATLFLPILASMSRSIEVNPFLYIMIPCTLSTSFAPMLPVA 528  
Qy 505 TPNALVAGYANIRTKDMAIAGIGPIITITLTFVFCQWGLVVPVNLNSFPPEWA 559  
Db 529 TPNNAIVFSYGHRLVSDVMVKTGIVMNIIGIITTSINTWGRPMF-SLDTFPQMA 582

RESULT 11  
S132\_HUMAN STANDARD; PRT; 592 AA.  
ID S132\_HUMAN  
AC Q13183;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Solute carrier family 13, member 2 (Renal sodium/dicarboxylate  
DE cotransporter) (Na(+)/dicarboxylate cotransporter 1) (NaDC-1).  
GN Name=SLC13A2; Synonyms=NADC1, SDC11;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=96199379; PubMed=8967342;  
RA Fajor A.M.;  
RT "Molecular cloning and functional expression of a sodium-dicarboxylate  
RT cotransporter from human kidney.";  
RL Am. J. Physiol. 270:F642-F648 (1996).  
CC -I- FUNCTION: Cotransport of sodium ions and dicarboxylates such as  
CC succinate and citrate.  
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -I- SIMILARITY: Belongs to the SLC13A transporter family. NADC  
CC subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U26209; AAA98504.1; -;  
DR Genew; HGNC:10917; SLC13A2.  
DR MIM; 604148; -;  
DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
DR GO; GO:0005624; C:membrane fraction; TAS.  
DR GO; GO:0015361; F:low affinity sodium:dicarboxylate symporter. .; TAS.  
DR GO; GO:0006810; P:transport; TAS.  
DR InterPro; IPR001898; Na/sul\_sympot.  
DR Pfam; PF00939; Na\_sulph\_symp; 1.  
DR

```
DR TIGR00785; dass; 1.
DR PROSITE; PS01271; NA_SULFATE; 1.
KW Sodium transport; Symport; Transmembrane; Transport.
FT TRANSMEM 13 33 Potential.
FT TRANSMEM 53 73 Potential.
FT TRANSMEM 86 106 Potential.
FT TRANSMEM 114 134 Potential.
FT TRANSMEM 221 241 Potential.
FT TRANSMEM 274 294 Potential.
FT TRANSMEM 324 344 Potential.
FT TRANSMEM 371 391 Potential.
FT TRANSMEM 450 470 Potential.
FT TRANSMEM 482 502 Potential.
FT TRANSMEM 511 531 Potential.
FT TRANSMEM 545 565 Potential.
SQ SEQUENCE 592 AA; 64410 MW; 41137D6621A0872A CRC64;

Query Match
Best Local Similarity 34.3%; Score 1026; DB 1; Length 592;
Matches 219; Conservative 124; Mismatches 189; Indels 56; Gaps 10;

QY 26 LVVFLVPLCLPVMLLNEGAERFCMYLLVMAIFWVTEALPLYVTSMIPVAFPIMGMS 85
DB 15 LIVFFVPLILLPILVPSKEYACAYIILMALFWCTEALPLAVTALPFLILFPMGIVD 74
QY 86 SDQCRLYFKDTLVFMGGMVALAVEYCNLHKRLALRVQIVGCSRRRLHFLIMVTMF 145
DB 75 ASEVAVEYLKDSNLLFFGGLLVAIAVEHNLHKRLALRVLLVGVVPAPLILGFMVLTAF 134
QY 146 LSMWISNAACTAMCPITQAVLEELQAGVCKINHEPQVQIVGGNKKNEDEPPYTK-- 203
DB 135 LSMWISNTATSAMVPIAHAVLDLHSSQASSNVBE-----GNNPTFELQEPSPQREV 188
QY 204 -----ITLCYVLGIAYASSLGGCGIIGTATNLTKFG 235
DB 189 TKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCMSLVCVYSASIGGIATLTGTAPNLVJQG 248
QY 236 IYEAFKNSQTEQMDPFTFMFVSPMLVYTLITFVFLQWFMGL-WRPKSKAQSVQRGR 294
DB 249 QINSIFPQNGVNVFASFSPFAFTWVILLLAWLQILFLGFNFRKNFGGKQKQ--- 305
QY 295 EGADVAKKVIDQRYKDLGMSIHEIQVMILTFMVMVYFTRKPGIFLGMADLL--NSKD- 351
DB 306 EQQAAVCVIOTEHLLGPMWFAEKAISILFVILVLLWFTREPFGFLGMGNLAFNAXGE 365
QY 382 --IRNSMTIFVVMCMFLPANYAFRYCTRGGPV--PTGTPSLIITWKVFTQKVPNGL 407
DB 366 SMVSDGTVAIFGIITFIIPSKFGLTQDPENPGKLPGLG---LLDQKTVNQKMPWNI 421
QY 408 VFLGGGFALAGSKSGMAKIGNALIGLVLPNSVLLVLLVAVFLTAFSSNVAITAN 467
DB 422 VLLGGGFALAGSKSGSEWLNKLTPLQGVPAFAIILSLVATFECTSNVATT 481
QY 468 IIPVLAEMSLAIEIHPYLIPLPAGLACSMAPFLFVSTPPNALVAGYANIRTKOMAIAGI 527
DB 482 IFLPILASMAQAI CLHPYVMLPCTLATSLAPMLFVATPPNAIVPSFGDLKVLDMARAGF 541
QY 528 GPTIITITLTFVFCOTGLVVPVNLNSPPEVAQIYAAA-----ALGNKT 571
DB 542 LLNTIGVLIILAINSWGIPLF-SLHSPFSAQSNNTTAQCLPSLANTT 588

RESULT 12
Q642N5 PRELIMINARY; PRT; 624 AA.
AC Q642N5
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole body;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole body;
RA Klein S., Gerhardt D.S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC081299; AA031299.1; -.
KW Hypothetical protein.
SQ SEQUENCE 624 AA; 69245 MW; 7A518827C541A36E CRC64;

Query Match
Best Local Similarity 34.2%; Score 1024.5; DB 2; Length 624;
Matches 219; Conservative 123; Mismatches 194; Indels 79; Gaps 9;

QY 23 WKGL-----VFLVPLCLPVMLLNEGAERFCMYLLVMAIFWVTEALPLYVTSMIP 75
DB 5 WKILANRNYIFILVPLFLPLVPLVPTKEASCGFVILVMAIFWCTEALPLAVTALP 64
QY 76 VAFPIGMSSDQCRLYFKDTLVFMGGMVALAVEYCNLHKRLALRVQIVGCSRRRL 135
DB 65 LFFPMGIMDSTAVCSQYLKDTNMLFGLGLVALISVEKNLHKRLALRVLLVGVK 124
QY 136 HFLIMVTMFLSMWISNAACTAMCPITQAVLEELQAGVCKINHEPQVQIVG---- 188
DB 125 LLGFVVTAFLSMWISNTATTAMVPIAQAFLQQLHTSEERVDKPESEKKNVNGM 184
QY 189 -----GNKKNDEPPYTK----- 203
DB 185 YESVLPNGKMTLAVENTYATVNEGFEMQDSTKDPESKLEKESIGPIVIELEDEK 244
QY 204 -----ITLC--YVLGIAYASSLGGCGIIGTATNLTKGIYEAREFNKSTEQMDP 253
DB 245 QKLEKHLKICKMTLSVCYSASIGGIATLTGTPNLVMKGMDLPEPNNVNFASW 304
QY 254 MFYSVPSMLVYTLITFVFLQWFMGLWRPKSKAQSVQRGRGADVAKKVIDQRYK 313
DB 305 FGFAFTFVLLALSLSWLQPMFLGPNFKNGCGGSEKQEKRAFKIISGEHKLGA 364
QY 314 MSIHEIQVMILFIPVVMVYFTRKPGIFLGMADLLNSKD-----IRNSMTIFV 369
DB 365 MTFAEISVLVFLILLVWFTREPFGFPGWATISFNKDGKEMVTDATVAIVSLMMF 424
QY 370 ANYAFLEY-----CTRRGSPVPT-CPTPSLITWKVFTQKVPNGLVFLGGG 424
DB 425 SELPSFKYREDTDEKPGMKPKFVPPALLDQKTVNEKMPWNIIVLLGGGFALAG 484
QY 425 GMAKIGNALIGLVLPNSVLLVLLVAVFLTAFSSNVAITANIIIPVLAEMSLAIE 484
```

Db 485 GLSLWLGKLTPLQSIIPPAIAIALILCLLVATFECTSNVATTTFLFLPILASMAKAIHLNP 544

Qy 485 LYILPAGLACSAEHLVPSTPNALVAGVANIRTKDMATAGIGPTIITITILFVFCQTW 544

Db 545 LYIMLPCTLSASLAFMLPVATPPNATAFSGYQKLVDMAXAGLLNLGVITITILAINSW 604

Qy 545 GLVYVYNLNSFPPEWA 559

Db 605 GFYMF-DLGTFFPSWA 618

RESULT 13

S132 RABIT STANDARD; PRT; 593 AA.

AC Q28615;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Solute carrier family 13, member 2 (Renal sodium/dicarboxylate cotransporter) (Na(+)/dicarboxylate cotransporter 1) (NADC-1).

DE Names=SLC13A2; Synonyms=NADC1, SDCT1;

GN Oryctolagus cuniculus (Rabbit)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI\_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=New Zealand white; TISSUE=Kidney cortex;

RX MEDLINE=95197599; PubMed=7890707;

RA Pajor A.M.;

RT "Sequence and functional characterization of a renal sodium/dicarboxylate cotransporter";

RL J. Biol. Chem. 270:5779-5785(1995).

CC -1- FUNCTION: Cotransport of sodium ions and dicarboxylates such as succinate and citrate.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- TISSUE SPECIFICITY: Abundant in kidney and small intestine.

CC -1- SIMILARITY: Belongs to the SLC13A transporter family. NADC subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; U12186; AAA99666.1; -.

DR PIR; I46528; I46528.

DR InterPro; IPR001898; Na/sul\_sympot.

DR Pfam; PF00939; Na\_sulph\_symp; 1.

DR TIGRFAMs; TIGR00785; class; 1.

DR PROSITE; PS01271; NA\_SULFATE; 1.

KW Sodium transport; Sympot; Transmembrane; Transport.

FT TRANSMEM 11 Potential.

FT TRANSMEM 31 Potential.

FT TRANSMEM 53 73 Potential.

FT TRANSMEM 86 106 Potential.

FT TRANSMEM 121 141 Potential.

FT TRANSMEM 221 241 Potential.

FT TRANSMEM 270 290 Potential.

FT TRANSMEM 327 347 Potential.

FT TRANSMEM 369 389 Potential.

FT TRANSMEM 451 471 Potential.

FT TRANSMEM 485 505 Potential.

FT TRANSMEM 514 534 Potential.

FT TRANSMEM 543 563 Potential.

QY SEQUENCE 593 AA; 65125 MW; 29386DB915D9E2BC CRC64;

Query Match 34.28; Score 1024; DB 1; Length 593;

Best Local Similarity 37.66; Pred. No. 1.9e-68;

Matches 219; Conservative 127; Mismatches 186; Indels 50; Gaps 10;

QY 20 ANHWK-----LVVLPVLLCLPVLMLNEGAEFCMYLLLVMAIFPWTEALPLVVTSM 72

Db 2 ATCWQLWAYRYMVLVFLPISLLPLVLPVPRKEAYCAVAILMALFWCTDALPLAVTAL 61

QY 73 IPVAPPINGIMSSDOTCLRYFKDTLYMFMGGIYVWALAVEYCNLHRLALRVITQIVGCSP 132

Db 62 LPLCLFFPMGIMEASEVGLYKDTNVLFIGGLLAIARVHWNHLKRIALRVLLLTGVRP 121

QY 133 RLHFLGLIMVTFMFLSMWISNAACATAMCPPIIOAVLEELQ-----AQGVCKINHE----- 181

Db 122 ALLILGFVVTATFLSMWISNTASTAMVPIARVLOELNNTQSNVEGSDNPFLEQEPS 181

QY 182 PQYQIVGNGKNKQNEDEPPYPT-----KITLCYLYGIAYASSLGGCGT 223

Db 182 POKETSKVDEKONGQAOPLPAVPLESGEHWTOQLRFSQGMSLC-----VCYSASIGGIAT 237

QY 224 IGTATNLTFKIYEAPFKNSTEQMDPPTFMFVSVPMSLVYTLITFVFLQHHFMGLWRPK 283

Db 238 LTGTTPLNLVQOMTSLFPQNPVNVNFASFQGFAPFIMVILLLLSLWLQILFLGINFRK 297

QY 284 SKEAQEVQRREGADVAKKVIDORYKDLGPMSTHETQVMILFIEMVMVFTFRKPGIFLGW 343

Db 298 NGFIREQEHQQRKQAAVYVITQYELLGLGPMSPAFAKAVILFVILLVLLFTFRPGFHHG 357

QY 344 ADLNSKD-----IRNSMPTIFVVMVCMPLPANYAFRLRYCTRRGGVPVPTPSLITWKF 398

Db 358 GNLVFSDDASGRVMVSDGSASILIGVFLFVFPVKIPGLTQDPDNPGRLLKA--PPALLNNWKL 415

QY 399 IQTKVPWGLVFLIGGCFALAEKSGQSMAGLIGNALIGLK-VLPNSVLLLVLLVAVELT 457

Db 416 VNKQPMWNLVLLGGGYALAKGSEGLSGLWLNKMLPQHVPPATVFIICULVATF-T 474

QY 458 AFSSNVAIAIIPVLAEMSLAIEIHPYLLIIPAGLACSMAPHLVPSTPNALVAGVANI 517

Db 475 ECTSNAATITLLPILASMAQAICLHPVLMVLPCTLASSLAFMLPVATPNAIVSEFGL 534

QY 518 RTKDMAIAGIGPTIITITILFVFCQWGLVYVYNLNSFPPEWA 559

Db 535 RVSDMARAGIMLNIIIGVLVIMLAINSGVPMF-QLHTFPSWA 575

RESULT 14

Q803K7 PRELIMINARY; PRT; 613 AA.

AC Q803K7;

DT 01-JUN-2003 (T-EMBLrel. 24, Created)

DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE Similar to solute carrier family 13, member 2.

GN ORFNames=zgc:55601;

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

OX NCBI\_TaxID=7955;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AB; TISSUE=Whole body;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426303899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler S.J., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udell I.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB; TISSUE=Whole body;  
 RA Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC044437; AAH44437.1; -  
 DR ZFIN; ZDB-GENE-040426-2803; zgc:55601.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006814; P:sodium ion transport; IEA.  
 DR InterPro; IPR001898; Na/sul\_sympot.  
 DR Pfam; PF00939; Na\_sulph\_symp; 1.  
 SQ SEQUENCE 613 AA; 67594 MW; F6B73FE343F288C8 CRC64;  
 Query Match 34.2%; Score 1023.5; DB 2; Length 613;  
 Best Local Similarity 36.5%; Pred. No. 2.2e-68;  
 Matches 217; Conservative 122; Mismatches 192; Indels 63; Gaps 9;  
 QY 22 HWKGLVFLVPLLCPLPVMMLNEGAFRCMYLLVNAIFWTEALPLYVTSMIPVAPPIM 81  
 DB 16 HRCNCLILCVPLILPLPLVPTPEARGCFAILMALFECMPLAITALLPVVLPFPM 75  
 QY 82 GIMSSDQTCRLYFKDTLVFMFGIMVALAVEYCNLHKLRLAVIOVGCSPRLHFLGIM 141  
 DB 76 GIMSEGVCVQLKDTNMLFVGLLVAIVAHNHLKRLIALSVLLVGVRRPALLMLGPMI 135  
 QY 142 VTMLFSLMWSNAACATMCPITIOAVLEEL-----QAQGVCKINHEPQ 183  
 DB 136 VTAFSLMWSINATTATMMLPISQAVLEQLSATEADSEKELREGVQNAFELTEVNIKOP 195  
 QY 184 YQIVGCKNKNE-----DEPPYTKITLCYLGIAVASSLGGCGTIIG 226  
 DB 196 LDNTQGEKPNNAADVEGINALSERRRKAREAKY-LRLFKGMSLSVCYSASIGGTATLTG 254  
 QY 227 TATNLTFGIYEAREFNKSTEQDPTFMFYSVPSMLVTLTTFVLOWHFMGLWRPKSKE 286  
 DB 255 TTPNLILKGQDEIFPDNDNDVINFAWFGFAPPNNVLMVLSWLVLCQCYLGF---NFKQ 311  
 QY 287 AQEVQRREGADVAKVIDQRYKDLGPMSTHEIOVWILFIEMVVMYFTRKPGIFLGWDL 346  
 DB 312 SFGCGTKNEGDKDAYKVMKNEYKKGPMSPFAGAVLVIFILVILWFTREGPFMGWATE 371  
 QY 347 L---NSKDIRNSMPTIFVVMCFMLPANYAFURYCT-----RRGFPVPTG----- 388  
 DB 372 LFNKNGQVVTGTVAFINSTLFFVIPSVDL--CSIKYEEQDEBAVGEQDGVBEKKR 429  
 QY 399 ---PPSLITWKFQTKVPWGLVFLGGFALAEKSGKAGKLLGNALIGKVLPSNVL 445  
 DB 430 KLKGTPTLLNWKVHERPMFNWVLLGGGFALASGESGLSVWLGQSLSPKSPFPAI 489  
 QY 446 LLVTLVAVFLTAPSSNAIANIIPVLAENSLAIEIHPYLILPAGLACSNAPFLPVST 505  
 DB 490 SIILCLLVGTTECSNTATTLTFLPILASHMATTIGLHPLVLMCTTISASLAFMLPVAT 549  
 QY 506 PPNALVAGYANIRTKDMAIAGIPTITITILFVFCQTWGLVVPVNLNSFPPEWA 559  
 DB 550 PPNALFSGYGNLKVLDMAKAGFILNIIGILCINLINTWGMAMF-KLDTFFPSWA 602  
 RESULT 15  
 Q6P3L2 ID PRELIMINARY; PRT; 613 AA.  
 AC Q6P3L2;  
 DT 05-JUL-2004 (Tremblrel. 27, Created)  
 DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)  
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
 DE zgc:77607 protein.

GN ORFNames=zgc:77607;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T.I., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Tothiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RA Strausberg R.;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC063947; AAH63947.1; -  
 DR ZFIN; ZDB-GENE-040426-2389; zgc:77607.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006814; P:sodium ion transport; IEA.  
 DR InterPro; IPR001898; Na/sul\_sympot.  
 DR Pfam; PF00939; Na\_sulph\_symp; 1.  
 SQ SEQUENCE 613 AA; 67653 MW; 8A4B42760D510C6B CRC64;  
 Query Match 34.2%; Score 1022.5; DB 2; Length 613;  
 Best Local Similarity 36.5%; Pred. No. 2.6e-68;  
 Matches 217; Conservative 122; Mismatches 192; Indels 63; Gaps 9;  
 QY 22 HWKGLVFLVPLLCPLPVMMLNEGAFRCMYLLVNAIFWTEALPLYVTSMIPVAPPIM 81  
 DB 16 HRCNCLILCVPLILPLPLVPTPEARGCFAILMALFECMPLAITALLPVVLPFPM 75  
 QY 82 GIMSSDQTCRLYFKDTLVFMFGIMVALAVEYCNLHKLRLAVIOVGCSPRLHFLGIM 141  
 DB 76 GIMSEGVCVQLKDTNMLFVGLLVAIVAHNHLKRLIALSVLLVGVRRPALLMLGPMI 135  
 QY 142 VTMLFSLMWSNAACATMCPITIOAVLEEL-----QAQGVCKINHEPQ 183  
 DB 136 VTAFSLMWSINATTATMMLPISQAVLEQLSATEADSEKELREGVQNAFELTEVNIKOP 195  
 QY 184 YQIVGCKNKNE-----DEPPYTKITLCYLGIAVASSLGGCGTIIG 226  
 DB 196 LDNTQGEKPNNAADVEGINALSERRRKAREAKY-LRLFKGMSLSVCYSASIGGTATLTG 254  
 QY 227 TATNLTFGIYEAREFNKSTEQDPTFMFYSVPSMLVTLTTFVLOWHFMGLWRPKSKE 286  
 DB 255 TTPNLILKGQDEIFPDNDNDVINFAWFGFAPPNNVLMVLSWLVLCQCYLGF---NFKQ 311  
 QY 287 AQEVQRREGADVAKVIDQRYKDLGPMSTHEIOVWILFIEMVVMYFTRKPGIFLGWDL 346  
 DB 312 SFGCGTKNEGDKDAYKVMKNEYKKGPMSPFAGAVLVIFILVILWFTREGPFMGWATE 371  
 QY 347 L---NSKDIRNSMPTIFVVMCFMLPANYAFURYCT-----RRGFPVPTG----- 388

```

| | : : : | | : | | : | | : |
Db 372 LFNKNGQVTDGTVAIFMSTLFFVIPSVDFL--CSIKYEEQDEEAVGEQDQGEKEKKR 429
Qy 389 ---PTPSLITWKFIQTQKVPWGLYFLGGGFALAEQSKQSGMAKLIQNALIGLKVLPNSVL 445
Db 430 KLGKTPDLLNWKVHERMPWNI VLLGGGFALASGESEGLSVWLGQSUSPLKSI PPFAT 489
Qy 446 LLVVILVAVFLTAFSSNVAIANIIPVLAEMSLAIEIHPYLIPLAGLACSMAPHLFPVST 505
Db 490 SIILCLLVGTFTBCSSNTATTLFLPLILASMTTIGLHPLYVMLPCTTISASLAFMLPVAT 549
Qy 506 PPNALVAGYANIRTKMAIAGIGFTIITITLTFVFCOTWGLVYVYPNLNSPPEWA 559
Db 550 PPNAIAPSYGNLKYLDMAKAGFILNIIGILCINILGINTWGMAMF-KLDTPPSWA 602
```

Search completed: June 30, 2005, 08:58:41  
Job time : 82.9736 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:14 ; Search time 79.2935 Seconds  
(without alignments)  
2741.199 Million cell updates/sec

Title: US-10-017-479A-3  
Perfect score: 2897  
Sequence: 1 MABPGQRKFLVGRCCIFHW.....FPESKSPDWAKEIKNQTKI 562

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

- Database : A\_Geneseq\_16Dec04:\*
- 1: Geneseqp1980s:\*
  - 2: Geneseqp1990s:\*
  - 3: Geneseqp2000s:\*
  - 4: Geneseqp2001s:\*
  - 5: Geneseqp2002s:\*
  - 6: Geneseqp2003as:\*
  - 7: Geneseqp2003bs:\*
  - 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	2114.5	73.0	908	4	ABB61068 Drosophil
2	1382.5	47.7	572	4	ABB60315 Drosophil
3	1382.5	47.7	572	4	ABB66903 Drosophil
4	1382.5	47.7	572	5	ABB79611 Drosophil
5	1382.5	47.7	572	8	ADP64789 Drosophil
6	1374.5	47.4	572	6	ABR40099 dindya s
7	885	30.5	568	7	ADK51050 Human NOV
8	885	30.5	568	8	ADH42445 Human NOV
9	885	30.5	568	8	ADN61775 Human NOV
10	885	30.5	568	8	ADN61775 Human NOV
11	875	30.2	568	6	ABR57023 Human TCH
12	875	30.2	568	7	ADK38764 Human 696
13	875	30.2	568	7	ADK51052 Human NOV
14	875	30.2	568	8	ADH42441 Human NOV
15	875	30.2	568	8	ADP64793 Human Na+
16	875	30.2	568	8	ADH42443 Human NOV
17	874.5	30.2	568	7	ADK51052 Human NOV
18	871	30.1	568	6	ABR40097 Human sod
19	871	30.1	587	7	AAE38766 Rat 69624
20	868	30.0	616	5	ABU65064 Human NOV
21	868	30.0	616	5	ABU65063 Human NOV
22	868	30.0	616	7	ADK51048 Human NOV
23	868	30.0	616	8	ADH42447 Human NOV
24	868	30.0	616	8	ADN61777 Human NOV
25	868	30.0	616	8	ADN61779 Human NOV

26	866	29.9	568	6	ABG75835	Abg75835 Transport
27	863	29.8	568	5	AAU79946	Aau79946 Human tra
28	863	29.8	568	8	ADG16993	Adg16993 Human tra
29	856.5	29.6	572	8	ADP64791	Adp64791 Rat INDY.
30	852.5	29.4	619	8	ADG16995	Adg16995 African c
31	849.5	29.3	581	8	ADP64799	Adp64799 Zebrafish
32	844	29.1	592	6	ABB82950	Abb82950 Human SLC
33	844	29.1	592	6	ABB82951	Abb82951 Human SLC
34	844	29.1	592	6	ABR40100	AbR40100 Human ren
35	844	29.1	592	6	ABR57025	AbR57025 Human NaD
36	844	29.1	592	7	ADP63755	Adp63755 Human pro
37	836.5	28.9	539	5	AAE21181	Aae21181 Human TRI
38	833.5	28.8	531	6	ABG75826	Abg75826 Transport
39	831.5	28.7	602	6	ABB82952	Abb82952 Human SLC
40	831.5	28.7	602	6	ABR40101	AbR40101 Human ren
41	831.5	28.7	602	6	ABG75837	Abg75837 Transport
42	814.5	28.1	572	6	ABR57024	AbR57024 Mouse TCH
43	814.5	28.1	572	8	ADP64797	Adp64797 Mouse Na+
44	812.5	28.0	551	8	ADR09550	Adr09550 Human pro
45	743	25.6	626	3	AAAB36167	Aab36167 Novel hum
46	742.5	25.6	627	3	AAAB36161	Aab36161 Novel hum
47	742	25.6	622	5	ABU65065	Abu65065 Human NOV
48	742	25.6	622	7	ADK51054	AdK51054 Human NOV
49	742	25.6	622	8	ADH42449	Adh42449 Novel hum
50	742	25.6	622	8	ADN61781	Adn61781 Human nov
51	741	25.6	626	3	AAAB36164	Aab36164 Novel hum
52	740.5	25.6	627	3	AAAB23625	Aab23625 Human sec
53	740.5	25.6	627	3	AAAB36158	Aab36158 Novel hum
54	738.5	25.5	627	3	AAAB42213	Aab42213 Human ORF
55	738.5	25.5	627	5	ABR97450	AbR97450 Novel hum
56	737	25.4	557	8	ADP29692	Adp29692 Human sec
57	737	25.4	626	3	AAAB36168	Aab36168 Novel hum
58	736.5	25.4	627	3	AAAB36162	Aab36162 Novel hum
59	735	25.4	626	3	AAAB36165	Aab36165 Novel hum
60	734.5	25.4	627	3	AAAB36159	Aab36159 Novel hum
61	730.5	25.2	519	8	ABM84184	Abm84184 Human dia
62	721.5	24.9	519	8	ADH22543	Adh22543 Human tra
63	708	24.4	533	8	ABM84303	Abm84303 Human dia
64	698	24.1	520	7	ADM05177	Adm05177 Human pro
65	696	24.0	595	5	AAE22910	Aae22910 Human tra
66	696	24.0	595	5	AAO21807	Aao21807 Lung-spec
67	696	24.0	595	6	ABB82949	Abb82949 Human SLC
68	696	24.0	595	7	AD121045	Ad121045 Novel hum
69	696	24.0	595	8	ADO78130	Ado78130 Human SLC
70	693.5	23.9	516	5	ABU65066	Abu65066 Human NOV
71	693.5	23.9	516	7	ADK51056	AdK51056 Human NOV
72	693.5	23.9	516	8	ADH42451	Adh42451 Novel hum
73	679	23.4	477	8	ADN61783	Adn61783 Human nov
74	671	23.2	551	8	ADP64795	Adp64795 Nematode
75	624	21.5	580	3	AAAB36169	Aab36169 Novel hum
76	623.5	21.5	581	3	AAAB36163	Aab36163 Novel hum
77	622	21.5	580	3	AAAB36166	Aab36166 Novel hum
78	621.5	21.5	581	3	AAAB36160	Aab36160 Novel hum
79	601.5	20.8	552	6	AAO31005	Aao31005 Human tra
80	588	20.3	543	6	ADA48682	Ada48682 Rice prot
81	588	20.3	543	7	ADC08241	Adc08241 Rice prot
82	551	19.0	510	4	AAG90000	Aag90000 C glutami
83	549	19.0	432	6	ABM71452	Abm71452 Staphyloc
84	545	18.8	502	6	ADA48226	Ada48226 Rice prot
85	545	18.8	552	2	AAW98815	Aaw98815 H. pylori
86	544	18.8	527	4	AAW76767	Aaw76767 Corynebact
87	544	18.8	540	3	AAAG51055	Aag51055 Arabidops
88	544	18.8	557	3	AAAG51054	Aag51054 Arabidops
89	543	18.7	540	3	AAAG13103	Aag13103 Arabidops
90	543	18.7	558	3	AAAG13102	Aag13102 Arabidops
91	539.5	18.6	466	3	AAAG51056	Aag51056 Arabidops
92	538.5	18.6	466	3	AAAG13104	Aag13104 Arabidops
93	535	18.5	407	5	AAU91115	Aau91115 Human sec
94	535	18.5	407	5	ABG65198	Abg65198 Human alb
95	535	18.5	407	8	ADL78465	Adl78465 Albumin f
96	511.5	17.7	377	5	AAU91090	Aau91090 Human sec
97	511.5	17.7	377	5	ABG65200	Abg65200 Human alb
98	511.5	17.7	377	8	ADL78467	Adl78467 Albumin f

99 487.5 16.8 382 7 ADE09041 Novel pro

100 459.5 15.9 462 6 ABM70253 Photorhab

ALIGNMENTS

RESULT 1

ABB61068

ID ABB61068 standard; protein; 908 AA.

XX ABB61068;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 9996.

DE Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

FA Venter JC, Adams M, Li PWD, Myers EW;

PI WPI; 2001-656860/75.

DR N-PSDB; ABL05171.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell

PT interactions.

XX Disclosure; SEQ ID NO 9996; 21pp + Sequence Listing; English.

PS The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 908 AA;

XX Query Match 73.0%; Score 2114.5; DB 4; Length 908;

XX Best Local Similarity 49.8%; Pred. No. 2.1e-210;

XX Matches 452; Conservative 20; Mismatches 38; Indels 397; Gaps 2;

QY 44 MAEFKCLWLVITMALLTETLPIYVVTALFPLVFCPLGLVNASIVCKQYFTDITVIVFLG 103

DB 1 MAEFKCLWLVITMALLTETLPIYVVTALFPLVFCPLGLVNASIVCKQYFTDITVIVFLG 60

QY 104 GLIVALGIEYNLHTRIALRVIRIVGSPRRLFVGLMSVSTFMGLWISNSAGTAMMCPIV 163

DB 61 GLIVALGIEYNLHTRIALRVIRIVGSPRRLFVGLMSVSTFMGLWISNSAGTAMMCPIV 120

QY 164 KALVNELDTNKIFPVYVMTQEEPEEGEPHPSKITVAFYAGIAYASSIGGLTIGTGT 223

DB 121 KALVNELDTNKIFPVYVMTQEEPEEGEPHPSKITVAFYAGIAYASSIGGLTIGTGT 180

QY 224 NLVFRGIYTERPPTSTVEITFANMFYSIPLMVI VNVTLVIIAFILITHMGLFRPNRSKTKG 283

DB 181 NLVFRGIYTERPPTSTVEITFANMFYSIPLMVI VNVTLVIIAFILITHMGLFRPNRSKTKG 240

QY 284 ITAEANTNRKLMEDVLRQRHIDLGPMSCHEIQMAIAFAPMIVLLITRKPGFVPGWSDLIN 343

DB 241 ITAEANTNRKLMEDVLRQRHIDLGPMSCHEIQMAIAFAPMIVLLITRKPGFVPGWSDLIN 300

QY 344 RKVGSASGLSFIVLLIFALPTQYTFYKCCGKGPFTAQDAIDAILSWEVVLNIPWGLLF 403

DB 301 RKVGSASGLSFIVLLIFALPTQYTFYKCCGKGPFTAQDAIDAILSWEVVLNIPWGLLF 360

QY 404 LGGGFALAVASRETGLNIMISKAMQVLIGLNPVQVSIT 443

DB 361 LGGGFALAVASRETGLNIMISKAMQVLIGLNPVQVSITADGKIGGGQAVKRCGCFHWR 420

QY 444 443

DB 421 GKATAMIPLITLPIMIYGVLENLNASYGVVDILESTSEKVCSEYFSDTVVMFIGLLIAL 480

QY 444 443

DB 481 AIEYSNLHORIALNTILIVGCSPRRLHFGLVMTFCISLWISNSAATAMCPIVKAVLNE 540

QY 444 443

DB 541 METQNIFAIYKTOBEEPVVEEGDPPHPSTISMAYFGIAYSSIGGCGTLTGCTNLTYKG 600

QY 444 443

DB 601 LYDTRPPNSDEKIDPPIFMAYSPFVVLVILFTVPSLQVTHMGLFRPNRSKIQEVKKA 660

QY 444 443

DB 661 ESQDVVQVIVKORKAELGPMSCHEIQVGLLFVLMIFLLETRKPGPPGHWADFLNKAIGS 720

QY 444 448

DB 721 GPPVFFATILLPALPTQYTFYKCCGKAPPGQTLDACLSWGFALAEGRSVSGMAKMLGE 780

QY 449 486

DB 781 SLAFAGEMHSVLVISMCIISLFCFASNAVAICNLIPIFSEMALAIEVHPMKLTFPPAA 840

QY 487 LGISMYVFLPVSTPPNAIVTQVAHIKTYFACCGIVPTIIGISVALVNTNMTWGLIIPES 546

DB 841 LACSLAFHLIPVSTPPNAILISGFTGLTKYMAIAGILPTCWAFLCLLFTGTGVTMTLIVPGT 900

QY 547 KSFPDWA 553

DB 901 TEPFPSWA 907

RESULT 2

ABB60315

ID ABB60315 standard; protein; 572 AA.

XX ABB60315;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 7737.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.



DB 240 RFKNSTQMDPTFMFVSFMSLV--YTLTFLVFLQWFMGLWRPKSKEAQEVQRGREGA 297  
 QY 293 KLMEVLRQRHIDLGPMSCHEIQMAIAFAPMIVLLITRKPGFVPGWSDLNKRVGSASG 352  
 DB 298 DVAKKVIDQYKDLGPMSCHEIQMAIAFAPMIVLLITRKPGFVPGWSDLNKRVGSASG 357  
 QY 353 LSFIVLLIFALPQYTFKFCCKG-GPPTAQADAILSWYVLRNIPWGLLFLGGSFAL 411  
 DB 358 TIFVVMCFMPLPANYAFRLYCTRRGGVPTGPTSLITWKFIQTKVPWGLVFLGGSFAL 417  
 QY 412 AVASRETNLMISKAMQVLIGLNPVIVQSIITFVLNFPFSAFNANVAVNIVLPILCEMS 471  
 DB 418 AEGSKQSGMAKLGKLVLPNSVLLVILVAVLTFAPSSNAIAIIPVLAEMS 477  
 QY 472 LALEHPLILTLPACLGISMVYFPLVSTPPNAIVTOYAHIKTKYFACCGIVPTIIGISVA 531  
 DB 478 LAIEIHPVLILPAGLACSMFHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITIIL 537  
 QY 532 LVNTNTWGLIIFPESKSFDPWAK 554  
 DB 538 FVFCQTWGLVVPNLNSFPPEWAQ 560

RESULT 4  
 ABB79611 standard; protein; 572 AA.

XX AC ABB79611;  
 XX DB 21-OCT-2002 (first entry)  
 XX DE Drosophila INDY protein dicarboxylate transporter.  
 XX KW Indy; dicarboxylate transporter; life span; longevity; obesity;  
 XX KW anorectic; caloric restriction; transgenic animal; gene therapy.  
 XX OS Drosophila melanogaster.  
 XX FN WO200259310-A2.  
 XX PD 01-AUG-2002.  
 XX PF 12-DEC-2001; 2001WO-US048130.  
 XX PR 12-DEC-2000; 2000US-0255013P.  
 XX PA (UYCO-) UNIV CONNECTICUT.  
 XX PI Rogina B, Reenan RA, Helfand SL;  
 XX DR WPI; 2002-599787/64.  
 XX DR N-PSDB; ABB84439.  
 PT IndY polynucleotide, useful for diagnosing or treating body weight  
 PT disorders, e.g. obesity, metabolic maintenance disorders, or the symptoms  
 PT of aging to extend the life span of an organism.  
 XX PS Claim 14; Page 81-83; 83pp; English.

CC The present sequence is the protein sequence of the Drosophila  
 CC melanogaster INDY protein, which is encoded by the IndY gene (see  
 CC ABB84439) involved in increased life span. The protein has similarity to  
 CC dicarboxylate transporters such as those from human and rat, and  
 CC represents a new class of dicarboxylate transporters that are not  
 CC inhibited by phloretin. Identification of the IndY gene resulted from the  
 CC observation that particular mutations in the gene caused an increase in  
 CC the life span of the fly carrying the mutation. As a result of this  
 CC finding, it is now possible to identify and/or isolate Drosophila lines  
 CC with longer life spans, as well as to identify agents that contribute to  
 CC longer life span. It is also possible to isolate genes involved in, and  
 CC which have an effect on longevity, as well as proteins encoded by these  
 CC genes. The invention provides IndY polynucleotides, proteins, anti-INDY

CC antibodies, antagonists that inhibit IndY activity or expression and  
 CC agonists that increase IndY activity or expression, and their use in the  
 CC diagnosis or treatment of body weight disorders, such as obesity and  
 CC metabolic maintenance disorders, or longevity in humans and animals.  
 CC Antagonists include at least a portion of the IndY gene sequence, an  
 CC antisense oligonucleotide, a ribozyme, a triple helix-forming molecule, a  
 CC double-stranded interfering RNA, an anti-IndY antibody, or a mixture of  
 CC these. Methods of calorically restricting an organism and of extending  
 CC the lifespan of an organism by administering the antagonist are claimed  
 XX  
 SQ Sequence 572 AA;

Query Match 47.7%; Score 1382.5; DB 5; Length 572;  
 Best Local Similarity 46.5%; Pred. No. 1.9e-134;  
 Matches 262; Conservative 120; Mismatches 164; Indels 17; Gaps 7;  
 QY 3 EFGQRKFLVGRCCITP---HWKGAIIIIPLITLILYIGFQTDMAEFKCLMLIVTMALL 59  
 DB 4 EIGEOPQPPV-KCSNPFANHWKGLVFLVFLCLPVMLL---NEGABFRCLMILLVMAIF 59  
 QY 60 WITETLPIVVTALFPLVFCPLGLLVNASIVCKOYFTDTIIVFLGLLIVAGIEYSNLHTR 119  
 DB 60 WYTEALPLYVTSMIPVIAFPPIGIMSSDQTCRLYFKDTLVPMFGGIMVALAVEYCNLHXR 119  
 QY 120 IALRVIRIVGSPRRLFVGLMSVTFMGLWISNSAGTAMMCFIVKALVNELDTNKIPPVY 179  
 DB 120 LALRVIIQVGCSPRRLHFLGLIMVTFMSWISNAACTAMMCPILQIQLVBELOAQGVCKIN 179  
 QY 180 MTQEEFV-----EKGEPHPSKITVAFYAGIYASSIGGLTGLTGTNLVFRGIYTE 233  
 DB 180 HEFQIVGVGNKKNNEDEPPYPTKITLCYVLGIYASSIGGCGGTIIIGTATNLTFKGIYE 239  
 QY 234 RPTSTVEITFANFMFYSIPLMVIVNVLIVIAFLITH-MGLFRPNSTKTKIIAEANTNR 292  
 DB 240 RFKNSTQMDPTFMFVSFMSLV--YTLTFLVFLQWFMGLWRPKSKEAQEVQRGREGA 297  
 QY 293 KLMEVLRQRHIDLGPMSCHEIQMAIAFAPMIVLLITRKPGFVPGWSDLNKRVGSASG 352  
 DB 298 DVAKKVIDQYKDLGPMSCHEIQMAIAFAPMIVLLITRKPGFVPGWSDLNKRVGSASG 357  
 QY 353 LSFIVLLIFALPQYTFKFCCKG-GPPTAQADAILSWYVLRNIPWGLLFLGGSFAL 411  
 DB 358 TIFVVMCFMPLPANYAFRLYCTRRGGVPTGPTSLITWKFIQTKVPWGLVFLGGSFAL 417  
 QY 412 AVASRETNLMISKAMQVLIGLNPVIVQSIITFVLNFPFSAFNANVAVNIVLPILCEMS 471  
 DB 418 AEGSKQSGMAKLGKLVLPNSVLLVILVAVLTFAPSSNAIAIIPVLAEMS 477  
 QY 472 LALEHPLILTLPACLGISMVYFPLVSTPPNAIVTOYAHIKTKYFACCGIVPTIIGISVA 531  
 DB 478 LAIEIHPVLILPAGLACSMFHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITIIL 537  
 QY 532 LVNTNTWGLIIFPESKSFDPWAK 554  
 DB 538 FVFCQTWGLVVPNLNSFPPEWAQ 560

RESULT 5  
 ADP64789  
 ID ADP64789 standard; protein; 572 AA.

XX AC ADP64789;  
 XX DB 26-AUG-2004 (first entry)  
 XX DE Drosophila INDY, an Na<sup>+</sup>-coupled citrate transporter protein.  
 XX KW sodium-coupled citrate transporter; transmembrane citrate transporter;  
 XX KW lifespan; weight reduction; weight gain prevention; blood cholesterol;  
 XX KW triglyceride; low density lipopolsaccharide; glucose; obesity;  
 XX KW hyperlipidemia; hypercholesterolemia; INDY protein.  
 XX OS Drosophila melanogaster.



Db 180 HEPOYQIVGNGKKNNEDEPPYPTKITLCYYLGIAYASSLGGCGTIIIGTATNLTFRKGIYEA 239  
Qy 234 RFTSTVEITFANPMFYSIPLMIVNVTLLIIFLITH-MGLRPNKSKTKIITAENTNR 292  
Db 240 RPKNSTEQMDFFTFMFYSVPSMLV--YTLTTFVFLQWHFGLWRPKSKEAQEVQREGA 297  
Qy 293 KLMEVLQRHIDLPGMSCHIEIQMAIAFAFMIVILLITRKPGFVPGWSDLIINRKVVGSA 352  
Db 298 DVAKKVIDQYKDLGPNISIEIQMILFIWVMYFTRKPGIFGLWADLLNSKDIRNSMP 357  
Qy 353 LSFIVLIFALPQTYTFKCYCGK-GPFTAQADAILSWBYLNRNPWGLLFLGGGFAL 411  
Db 358 TIFVVMCFMLPANYAFRLYCTRGGGPVGTPTPSLITWKFIQTKVPWGLVFLGGGFAL 417  
Qy 412 AVASREGLNIMISKAMQVILGILPNIVVQSITFVLANFSAFNANVNVANIVLPILCEMS 471  
Db 418 AEGSKQSGMAKLIGNALIGLKVLPNSVLLLVILVAFLFAFSSNVAIANIIPVLAEMS 477  
Qy 472 LALEHPLILTPACLGISMVYFPLVSTPPNAIVTQVAHIKTKYFACCGIVPTIIGISVA 531  
Db 478 LAIEIHPYLLPAGLACSMFHLFVSTPPNALVAGYANIRTKDMAIGPITIIITL 537  
Qy 532 LVNTNTWGLIIPESKFPDPAK 554  
Db 538 FVFCQTWGLVVPNLNFPPEWAQ 560  
RESULT 7  
ABU65062  
ID ABU65062 standard; protein; 568 AA.  
XX AC ABU65062;  
XX DT 20-MAY-2003 (first entry)  
XX DE Human NOV14a protein.  
XX KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;  
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;  
KW human.  
XX OS Homo sapiens.  
XX PN WO200272757-A2.  
XX PD 19-SEP-2002.  
XX PF 08-MAR-2002; 2002WO-US006908.  
XX PR 08-MAR-2001; 2001US-0274101P.  
PR 08-MAR-2001; 2001US-0274194P.  
PR 08-MAR-2001; 2001US-0274281P.  
PR 08-MAR-2001; 2001US-0274322P.  
PR 09-MAR-2001; 2001US-0274849P.  
PR 12-MAR-2001; 2001US-0275235P.  
PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276776P.  
PR 19-MAR-2001; 2001US-0276994P.  
PR 20-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 21-MAR-2001; 2001US-0277791P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278999P.  
PR 27-MAR-2001; 2001US-0279036P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0277338P.

PR 30-MAR-2001; 2001US-0279995P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0280900P.  
PR 04-APR-2001; 2001US-0281194P.  
PR 13-APR-2001; 2001US-0283675P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
PR 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
PR 18-OCT-2001; 2001US-0330380P.  
PR 31-OCT-2001; 2001US-0335301P.  
PR 14-NOV-2001; 2001US-0332172P.  
PR 14-NOV-2001; 2001US-0332271P.  
PR 14-NOV-2001; 2001US-0332272P.  
PR 14-NOV-2001; 2001US-0333184P.  
PR 14-NOV-2001; 2001US-0333272P.  
PR 21-NOV-2001; 2001US-0332094P.  
PR 03-DEC-2001; 2001US-0337426P.  
PR 03-DEC-2001; 2001US-0338092P.  
PR 04-DEC-2001; 2001US-0337185P.  
PR 03-JAN-2002; 2001US-0345705P.  
PR 07-MAR-2002; 2002US-00092900.  
XX (CURA-) CURAGEN CORP.  
XX PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;  
XX PI Zerkhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;  
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;  
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;  
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;  
PI Lepley DM, Rieger DK;  
XX WPI; 2002-723332/78.  
XX N-PSDB; ABX97029.  
XX NOVX polypeptides and polynucleotides, useful for preventing or treating  
PT a disorder associated with aberrant NOVX expression or activity e.g.,  
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
PT asthma.  
XX Claim 1; Page 133; 1103pp; English.  
XX This invention describes novel human NOVX polypeptides which have  
XX cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive  
XX activity. Pharmaceutical compositions comprising the NOVX proteins or  
XX nucleic acid molecules or NOVX antibodies are useful for preventing or  
XX treating a disorder associated with aberrant NOVX expression or activity  
XX e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
XX asthma. The products of the invention can be used for gene therapy or in  
XX a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by  
XX ABX97008-ABX97185  
SQ Sequence 568 AA;  
Query Match 30.5%; Score 885; DB 5; Length 568;

PR	04-NOV-2002; 2002US-00287226.
PR	31-MAR-2003; 2003US-00403161.
XX	(CURA-) CURAGEN CORP.
PA	Anderson DW, Bento P, Boldog FL, Burgess CE, Caaman SJ, Furtak K;
XX	Gorman L, Gould-Rothberg BE, Gunther E, Heyes MP, Li L, Spytek KA;
PI	Stone DJ, Zhong M, Malyankar UM, Edinger SR, Patturajan M;
PI	Rothenberg ME, Smithson G;
XX	WPI: 2003-812539/76.
DR	N-PSDB; ADK51049.
DR	
XX	
PT	New NOVX polypeptide, useful for preparing a composition for treating or
PT	preventing e.g. cancer or for chromosome mapping.
XX	
PS	Claim 1; SEQ ID NO 70; 433pp; English.
XX	
CC	This invention relates to novel isolated polypeptides and the DNA
CC	sequences which encode them. The invention may be useful for the
CC	development of compounds with a cytostatic activity (as NOVX-agonists or
CC	antagonists) or vaccines. In addition, the disclosed sequences may be
CC	useful for gene therapy. The polypeptide is useful for preparing a
CC	composition for treating or preventing a pathological state in a mammal,
CC	for example cancer or for chromosome mapping. The present sequence is
CC	that of a human NOVX protein of the invention.
XX	
SQ	Sequence 568 AA;
	Query Match 30.5%; Score 885; DB 7; Length 568;
	Best Local Similarity 35.6%; Pred. No. 1.2e-82;
	Matches 197; Conservative 101; Mismatches 219; Indels 36; Gaps 9;
QY	27 IIPILITPLIYGFQDMAEFKCLMLIVTMALLMTITPLIYVVTALFPLVFCPLGLVNA 86
Db	19 VTPLLLPLVI-----LMPAKVSCAVIILMAIYVCTEVIPLAVTSLMPVLLPFLQILDS 74
QY	87 SVVCKQYFDTIVVFLGGLIVALGIEYSNLHRIALRVIRVGGSPRRLFVGLMSVSTFM 146
Db	75 ROVCVQYMKDTNMLFLGGLIYAVAVERNLHKRIALRTLLMWGAKPARMLGFMGTALL 134
QY	147 GLWISNAGTAMCPIVKNVNELDTN-----KIFPVYMTQEEBP-VEE 189
Db	135 SMWISNTATTAMVPIVEAILQOMEATSAATAGLELVDKGAKELPGSQVIFEGPTLQ 194
QY	190 GEPPHPSKITVAFYAGIAYASSIGGLTIGTGNLFRGIYTERFPTSTVTEITFANFMF 249
Db	195 QEDQERKRLCKAWTLCICVAASIGGTATLTGTGNVLLGQNNELFPDSKDLVNFASFPA 254
QY	250 YSIPLMVIVNVTLVIIAFL-ITHMGLFRPNSKTGKIIAEANTNRKLMEDVLRQRHIDLGP 308
Db	255 FAFPNMLV-----MLLFAMLMLOFYVMFSSFKKSGCGLESKNEKAALKVLOEYRKLGP 310
QY	309 MSCHEIQWAIAPAFMIVLLITRKPGFVPGWSDLI-----NRKVGSASGLSFVILLIFALP 364
Db	311 LSFABINVLICFPLLVILWFSRDPGFMGWLTVAVWEGETKYVSDATVAIFVATLLFIVP 370
QY	365 TOYTFKVC-----GKGPFTAQDAIILSWEYVURNIPWGLLFLGGGFALAVASRETLG 420
Db	371 SQPKFNFQSRQTEGKSPVLI-APPPLLDWKVTQEKVPWGVLLGGGFALAKGSEASGL 429
QY	421 NIMISKAMQVLIGLPNIVVQSITFVLNFFSAFNANVVANIVLPILCSEMSLALSLHPLI 480
Db	430 SVWGMKQMEPLHAVPPAAITLILSLVAVFECTSNVATTLFLPIFASMSRSIGLNPLY 489
QY	481 LTPACLGISMVYFLPVSTPPNAIVTOYAHIKTKYFACCGIVPTTIIGISVALVNTWGL 540
Db	490 IMPLPCTLSASFAMLPVATPPNAIVFTYGHKLVADMVKTVGIMNIIGVFCVFLAVNTWGR 549
QY	541 IIFPESKSPFDWA 553
Db	550 AIF-DLDHFPDWA 561
	RESULT 8
ID	ADK51050
XX	ADK51050 standard; protein; 568 AA.
XX	AC
XX	ADK51050;
DT	17-JUN-2004 (first entry)
DE	Human NOV18B protein sequence SeqID70.
XX	
KW	cytostatic; NOVX-agonist; NOVX-antagonist; vaccine; gene therapy; cancer;
XX	chromosome mapping; human; NOV18B.
OS	Homo sapiens.
XX	
PN	WO2003083046-A2.
XX	
PD	09-OCT-2003.
XX	
PF	01-APR-2003; 2003WO-US010142.
XX	
PR	02-APR-2002; 2002US-00115479.
PR	05-APR-2002; 2002US-0370349P.
PR	08-APR-2002; 2002US-0370969P.
PR	12-APR-2002; 2002US-0372019P.
PR	22-APR-2002; 2002US-0374379P.
PR	30-MAY-2002; 2002US-0384543P.
PR	03-JUN-2002; 2002US-00160619.
PR	15-AUG-2002; 2002US-0403748P.





PA (MALY/) MALYANKAR U M.  
PA (GERL/) GERLACH V.  
PA (LIUY/) LIU Y.  
PA (ANDE/) ANDERSON D W.  
PA (SPAD/) SPADERNA S K.  
PA (CATT/) CATTERTON E.  
PA (LEIT/) LEITE M W.  
PA (ZHON/) ZHONG H.  
PA (ALSO/) ALSOBROOK J P.  
PA (LEPL/) LEPLEY D M.  
PA (RIEG/) RIEGER D K.  
PA (BURG/) BURGESS C E.  
XX  
PI Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;  
PI Zerrhusen BD, Gusev VI, Ji W, Gorman L, Miller CE, Kekuda R;  
PI Patturajan M, Gangolli EA, Vernet CM, Guo XS, Tchernev VT;  
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;  
PI Anderson DW, Spaderna SK, Catterton E, Leite MW, Zhong H;  
PI Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;  
XX  
XX WPI; 2004-225693/21.  
XX N-PSDB; ADN61774.  
XX  
XX New NOVX polypeptides and nucleic acid molecules useful for diagnosing,  
XX preventing or treating NOVX-associated disorders, e.g. cancer, diabetes,  
XX infection or obesity, and in chromosome mapping, tissue typing or  
XX pharmacogenomics.  
XX  
XX Claim 1; SEQ ID NO 44; 786pp; English.  
XX  
XX The invention relates to an isolated polypeptide (designated NOVX, or  
XX NOV1-NOV127) comprising a sequence selected from 178 fully defined amino  
XX acid sequences (and their mature forms, variants and fragments). Also  
XX included are an isolated nucleic acid molecule encoding NOVX, a vector  
XX comprising the nucleic acid, a cell comprising the vector, methods for  
XX determining the presence or amount of the polypeptide or the nucleic acid  
XX molecule in a sample, methods for determining the presence of or  
XX predisposition to a disease associated with altered levels of expression  
XX of the above polypeptide or nucleic acid molecule in a first mammalian  
XX subject, a method for identifying an agent that binds to the above  
XX polypeptide, a method for identifying a potential therapeutic agent for  
XX use in the treatment of a pathology that is related to aberrant  
XX expression or physiological interactions of the polypeptide, a method of  
XX screening for a modulator of activity or of latency or predisposition to  
XX a pathology associated with the polypeptide and a method for modulating  
XX the activity of the polypeptide cited above. The composition and methods  
XX are useful for diagnosing, preventing or treating diseases such as  
XX diabetes, obesity, infectious diseases, anorexia, cancer-associated  
XX cachexia, cancer, neurodegenerative disorders like Alzheimer's disease or  
XX Parkinson's disease, immune disorders, haematopoietic disorders,  
XX dyslipidaemias, and other chronic diseases. These may also be used in  
XX chromosome mapping, tissue typing, preventive medicine and  
XX pharmacogenomics. The polypeptides are also useful as vaccines. The  
XX present sequence represents a NOVX protein of the invention.  
XX  
XX Sequence 568 AA;  
XX  
XX Query Match 30.5%; Score 885; DB 8; Length 568;  
XX Best Local Similarity 35.6%; Pred. No. 1.2e-82;  
XX Matches 197; Conservative 101; Mismatches 219; Indels 36; Gaps 9;  
XX  
Qy 27 IIPILYIGFTDMABPKLWLVITMALLWITITPLVYVTLPLVFCPLGLVNA 86  
Db 19 VTPULLPLVJ- ----LMPARKVCAYVILMAIYVCTEVIPLVTSIMPLVFLPQLDS 74  
Qy 87 SIVCKQYFTDITVIVFLGLIYVALGIEYSNLHRIALVRIVGSPRLFLVGLMSVSTFM 146  
Db 75 RQVCVQYMKDITNMLFLGLIYVAVVERWNLHRIALRTLWVGAKPARLMGLFGVTALL 134  
Qy 147 GLWISNAGTAMCPVIVKALVNEIDTN-----KIFPYMTQEEEP-VEE 189  
Db 135 SMWISNTATTAMVPVIVEAILQOMEATSAATEAGLELVDKGAKELPGSQVIFEGTLQG 194

QY 190 GEPPHPSKITVAFYAGIAYVASSIGGLTIGTGNLFRGIYTERPPTSTVETITFANFMF 249  
Db 195 QEDQERKRLCKAWTLICVYASIGGTATLTGTGNVLLGQNNELFPDPSKDLVNFASWEA 254  
QY 250 YSIPLMVIVNVTLVIAFL-ITHMGLFRNSKTGKIIAEANTNRKLMEDVLRQRHIDLGP 308  
Db 255 FAFPPNMLV- ---MLLPFWLWLFQVYMFSSFKSGWCGLESKKNEKAALKVLOEYRKLGP 310  
QY 309 MSCHETQMAIAFAFMIVLLITRKPFGVPCWSDLI- ---NRKVVGSAAGLSFVLLIFALP 364  
Db 311 LSAEINVLICFLLVILWFSRDPGFMGWLIVAVVEGETKVSVDATVAIFVATLLFIVP 370  
QY 365 TQYTFKYCC- ---GKGPFQAQIDAILSWEYVLRNIPWGLLFLGGGFALAVASRETGL 420  
Db 371 SQKRFNFRSQTSEEGKSPVLI-APPPELLDWKVQEKVPMGIVLLLGGGFALAGSEASGL 429  
QY 421 NMISKAMOVILGLNIVVQSTIFULANFFSFAFNANVVANIVLPILCMSLALEHPLI 480  
Db 430 SYVMGKQMEPLHAVPPAAITLILSLVAVFTECTSNVATTTILFLPIFASMSRSIGLNPLY 489  
QY 481 LTLPAICIGISMYVFLPVSTPPNAIVTOYAHITKYFACCGIVPTIIGISVALVNTNTWGL 540  
Db 490 IMLPCTLSASFAPMLPVATPPNAIVFTYCHLKVADMKVGTGVMNIIIGVFCVFLAVNTWGR 549  
QY 541 IIFPESKSPFDWA 553  
Db 550 AIF-DLDHFPDWA 561  
XX  
XX RESULT 11  
XX ABR57023  
XX ID ABR57023 standard; protein; 568 AA.  
XX AC ABR57023;  
XX XX 05-AUG-2003 (first entry)  
XX DE Human TCH169 protein SEQ ID NO:1.  
XX KW Human; TCH169; dicarboxylate transport; hepatotropic; cytostatic;  
XX nephrotropic; vasotropic; antidiabetic; liver disease; hepatitis;  
XX hepatic sclerosis; alcohol-related liver disease; prostate disease;  
XX prostatitis; prostatic hypertrophy; spleen disease; spleen hyperactivity;  
XX kidney disease; nephritis; kidney failure; nephritis; dropsy; diabetes;  
XX diabetes-associated renal disease; metabolic disease; hyperlipaemia;  
XX circulatory disease; arteriosclerosis; cancer.  
XX OS Homo sapiens.  
XX PN WO2003025168-A1.  
XX PD 27-MAR-2003.  
XX PF 13-SEP-2002; 2002WO-JP009444.  
XX PR 17-SEP-2001; 2001JP-00281992.  
XX PR 02-OCT-2001; 2001JP-00306873.  
XX PR 16-APR-2002; 2002JP-00113279.  
XX XX (TAKE ) TAKEDA CHEM IND LTD.  
XX PA Nakanishi A, Uno Y, Sagiya Y;  
XX WPI; 2003-313352/30.  
XX DR N-PSDB; ACC79539, ACC79558.  
XX XX Protein TCH169 with dicarboxylate transport activity for treatment and  
XX PT diagnosis of diseases including liver disease, cancer and circulatory  
XX PT disorders.  
XX XX Claim 1; Fig 1-2; 132pp; Japanese.  
XX PS The present invention describes protein TCH169 and its salts having  
XX CC



XX WPI: 2003-644620/61.  
DR N-PSDB; AAD58822.  
XX Novel 69624 polypeptide, a human transporter family member, useful for  
PT treating disorders e.g. hypocitraturia, such as hypocitraturia, formation  
PT of calcium stones, mental retardation abnormal body sulfate homeostasis.  
XX  
XX  
PS Claim 8; Page 43-45; Opp; English.  
XX  
XX The invention relates to 69624 polypeptide, a human transporter family  
CC member and its corresponding nucleic acid. 69624 protein is useful for  
CC developing novel diagnostic and therapeutic agents for 69624-mediated or  
CC related disorders. 69624 protein act as therapeutic or diagnostic agents  
CC for renal, neurological, colon or hepatic disorders. It act as diagnostic  
CC targets and therapeutic agents for treating disorders such as  
CC hypocitraturia, formation of calcium stones, mental retardation (Canavan  
CC disease) or abnormal body sulphate homeostasis. 69624 protein may act as  
CC diagnostic targets and therapeutic agents for controlling cellular  
CC proliferative and/or differentiative disorders such as carcinoma,  
CC sarcoma, metastatic disorder or haematopoietic neoplastic disorders e.g.,  
CC leukemia, immune disorders such as autoimmune disorders (diabetes  
CC mellitus, arthritis), multiple sclerosis, encephalomyelitis, myasthenia  
CC gravis, systemic lupus erythematosus, cardiovascular disorders such as  
CC arteriosclerosis, atherosclerosis, ischaemia reperfusion injury, cardiac  
CC hypertrophy, atrial fibrillation etc; and disorders involving abnormal or  
CC excessive pain. 69624 sequence is useful as pharmacodynamic marker and is  
CC also used in forensic identification of a biological sample. The present  
CC sequence is human 69624 protein  
XX  
SQ Sequence 568 AA;  
  
Query Match 30.2%; Score 875; DB 7; Length 568;  
Best Local Similarity 35.7%; Pred. No. 1.3e-81;  
Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;  
  
QY 27 IIPILITPILIIYGFQTDMAEF-KCLMLIVTMALLWITETPIYVTFALFPLVFCPLGLVN 85  
DB 19 VTPLLPLPLVI-----LMPAKFVRCAYIILMAYIWCTEVIPLAVTSLMPVLLPFLQILD 74  
QY 86 ASIVCKQYFTDTTIVFLGGLIIVAGIYBSNLHRIALRVIRIVGGSPRRLFVGLMSVSTF 145  
DB 75 SRQVCVQYKMDTNLFLTGLIIVAVAVERNLHRIALRTLLWVGAKPARMLGFMGVITAL 134  
QY 146 MGLWISNSAGTAMWCPTVKALVNELDN-----KTPPVYMTQBERP-VE 188  
DB 135 LSWMISNTATTAMVPVIEAILQOMETSATETAGLELVNDKAKELPGSQVIFEGPTIG 194  
QY 189 EGEPPHPKSTITVAFYAGIAYASSIGGLTIGTGNLVFRGIYTERPPTSTVITTFANFM 248  
DB 195 QQEDQERKRLCKAWTLCTCYAASIGGTATLTGTGPNVLLGQNNELFPDSDKLVNFASF 254  
QY 249 FYSIPLMIVNVVLTIIAFILITMGLRPN-SKTGKIIAANTRKLMEDVILQRHIDLG 307  
DB 255 AFAPPNMLV-----MLLFAWLWLPQVYMRFPFKKSWGCGLESKKNEKAALKVQLQEEYRKL 310  
QY 308 PMSCHETQMAIAFAMTVLLITRKPVPQGSDLI-----NRKVGSASGLSFTVLIFAL 363  
DB 311 PLFAEINVLICFLLIIVLWFSRDPGPNWGLTWAVWEGETKYVSDATVAIFATLLFIV 370  
QY 364 PTQYTFKYCC-----QKGPFPTAQADAILSWEYVLRNIPWGLLFLIGGGFALAVASRET 418  
DB 371 PSQPKFNFRSQTBEERKTFYP-----PPLDWKVTQKVPWGVILLGGGFALAKGSEAS 427  
QY 419 GLNIMISKANQVLICGLNIVVQSTTFVLNFFSFANANVVANIVLPILCMSIALLELHP 478  
DB 428 GLSYWVGKQEPLEHAPVPPAAITLLSLVAVFTECTSNVATTLFLFIPFASMSRSIGLNP 487  
QY 479 LILFLPACLGISMVYFLPVSTPPNNAIVTQVAHIKTKYFACCGIVPTTIGISVALVNTNTW 538  
DB 488 LYIMLPCTLSAFAPMLPVATPNAIVFTYGHKLVADWVKVTGIMNIIIGVFCVFLAVNTW 547  
QY 539 GLIIFPESKGFDDWA 553

DB 548 GRAIF-DLDHPPDWA 561  
RESULT 13  
ADKS1052  
ID ADKS1052 standard; protein; 568 AA.  
XX  
XX AC ADKS1052;  
XX DT 17-JUN-2004 (first entry)  
XX Human NOV18C protein sequence SeqID72.  
XX cytotstatic; NOVX-agonist; NOVX-antagonist; vaccine; gene therapy; cancer;  
XX chromosome mapping; human; NOV18C.  
XX Homo sapiens.  
XX WO2003083046-A2.  
XX 09-OCT-2003.  
XX 01-APR-2003; 2003WO-US010142.  
XX 02-APR-2002; 2002US-00115479.  
XX 05-APR-2002; 2002US-0370349P.  
XX 08-APR-2002; 2002US-0370969P.  
XX 12-APR-2002; 2002US-0372019P.  
XX 22-APR-2002; 2002US-0374379P.  
XX 30-MAY-2002; 2002US-0384543P.  
XX 03-JUN-2002; 2002US-00160619.  
XX 15-AUG-2002; 2002US-0403748P.  
XX 04-NOV-2002; 2002US-00287226.  
XX 31-MAR-2003; 2003US-00403161.  
XX (CURA-) CURAGEN CORP.  
XX Anderson DW, Bento P, Boldog FL, Burgess CE, Casman SJ, Furtak K;  
XX Gorman LJ, Gould-Rothberg BE, Gunther E, Heyes MP, Li L, Spytek KA;  
XX Stone DJ, Zhong M, Malyankar UM, Edinger SR, Patturajan M;  
XX Rothenberg ME, Smithson G;  
XX WPI; 2003-812539/76.  
XX N-PSDB; ADKS1051.  
XX New NOVX polypeptide, useful for preparing a composition for treating or  
XX preventing e.g. cancer or for chromosome mapping.  
XX Claim 1; SEQ ID NO 72; 433pp; English.  
XX This invention relates to novel isolated polypeptides and the DNA  
XX sequences which encode them. The invention may be useful for the  
XX development of compounds with a cytostatic activity (as NOVX-agonists or  
XX antagonists) or vaccines. In addition, the disclosed sequences may be  
XX useful for gene therapy. The polypeptide is useful for preparing a  
XX composition for treating or preventing a pathological state in a mammal,  
XX for example cancer or for chromosome mapping. The present sequence is  
XX that of a human NOVX protein of the invention.  
XX  
SQ Sequence 568 AA;  
  
Query Match 30.2%; Score 875; DB 7; Length 568;  
Best Local Similarity 35.7%; Pred. No. 1.3e-81;  
Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;  
  
QY 27 IIPILITPILIIYGFQTDMAEF-KCLMLIVTMALLWITETPIYVTFALFPLVFCPLGLVN 85  
DB 19 VTPLLPLPLVI-----LMPAKFVRCAYIILMAYIWCTEVIPLAVTSLMPVLLPFLQILD 74  
QY 86 ASIVCKQYFTDTTIVFLGGLIIVAGIYBSNLHRIALRVIRIVGGSPRRLFVGLMSVSTF 145  
DB 75 SRQVCVQYKMDTNLFLGGLIIVAVAVERNLHRIALRTLLWVGAKPARMLGFMGVITAL 134  
DB 75 SRQVCVQYKMDTNLFLGGLIIVAVAVERNLHRIALRTLLWVGAKPARMLGFMGVITAL 134



CC useful in diagnosing, treating or preventing NOVX-associated disorders,  
 CC e.g. cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,  
 CC diabetes, AIDS, multiple sclerosis, graft-versus-host disease,  
 CC Alzheimer's disease, Parkinson's disease, asthma, or fertility disorders.  
 CC The nucleic acids are further used as hybridization probes, in chromosome  
 CC mapping, tissue typing, preventive medicine, and pharmacogenomics. The  
 CC polypeptides are also useful as vaccines. This sequence represents an  
 CC example of the polypeptide of the invention.  
 XX  
 SQ Sequence 568 AA;

Query Match 30.2%; Score 875; DB 8; Length 568;  
 Best Local Similarity 35.7%; Pred. No. 1.3e-81;  
 Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;  
 QY 27 IIPILITPLIYGFQDQMAEF-KCLMLIVTMALLWITETPIYVITLFLVFCPLGLV 85  
 DB 19 VTPLLLPLVI-----LMPAKFVRCAYIILMAIYVCTEVIPLAVTSLMPVLLPQLD 74  
 QY 86 ASIVCKQYFDTIIVFLGGLIIVAGIEYSNLHTRIALRVIRIVGSGPRRLFVGLMSVSTF 145  
 DB 75 SRQVCVQYQKMDTNLFLGGLIIVAVVERNLHTRIALRVIRIVGSGPRRLFVGLMSVSTF 134  
 QY 146 MGLWISNSAGTAMCPVIVKALVNELDN-----KIPFVYMTQEEEP-VE 188  
 DB 135 LSMWISNTATTAMVPVIVEAILQOMEATSAATEAGLELVDKGAKEPLGSQVIFEGPTLG 194  
 QY 189 EGEPHPHSKITVAFYAGIAYASSIGGLTIGTGNLVFRGIYTERPPTSTVITFANFM 248  
 DB 195 QOEDQERKRLCKAMTLCICVAASIGGTATLTGTGPNVLLGQMNELFPDSDKLVNFASF 254  
 QY 249 FYSIPLMIVNVTVIITAFIITHMGLFRPN-SKTGKIIAEANTNRKLMEDVLQRHIDL 307  
 DB 255 AFAPFNNLV-----MLLPFWLWLFQVYVNRFNFKKSWGGGLSKKNEKALKVLQEEYRKL 310  
 QY 308 PMSCHEIQMAIAFAPMIVLLITRKPGFVPGWSDLI-----NRKVGSAGSLFIVLLIFAL 363  
 DB 311 PLPSFAEINVLICFLLVILWFSRDPGMPGWLTVAVWVEGETKVS DATVAIFVATLLFIV 370  
 QY 364 PTQYTFKYCC-----GKGPFTAQDAIILSWYVLRNIPWGLLFLGCGFALAVASRET 418  
 DB 371 PSQKPFNFPSQTEERKTPFPY---PPLDVKVTKQKVPWGVLLGGGFALAKSEAS 427  
 QY 419 GLNIMISKAMQVLIGLNPVQVSIITFVLANFFSAFNANVVVAVNIVLPILCEMSLALHLP 478  
 DB 428 GLSVWKGQKQEPHVAVPPAAITLILSLVAVFTCTSNVATTLFLFIPFASMSRSLG 487  
 QY 479 LILTPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTTIIGISVALVNTW 538  
 DB 488 LYIMLPCTLSASFAPMLPVATPPNAIVFTYGHKLVADMWKVTGVMIMNIIGVFCVFLAV 547  
 QY 539 GLIIPPSKSPDDWA 553  
 DB 548 GRAIF-DLDFPDWA 561

RESULT 15

ID ADP64793 standard; protein; 568 AA.  
 AC ADP64793;  
 XX  
 XX ADP64793;  
 DT 26-AUG-2004 (first entry)  
 DE Human Na<sup>+</sup>-coupled citrate transporter protein.  
 KW sodium-coupled citrate transporter; transmembrane citrate transporter;  
 KW lifespan; weight reduction; weight gain prevention; blood cholesterol;  
 KW triglyceride; low density lipopolysaccharide; glucose; obesity;  
 KW hyperlipidemia; hypercholesterolemia; INDY protein.  
 OS Homo sapiens.  
 XX

PN WO2004048925-A2.  
 XX 10-JUN-2004.  
 XX 20-NOV-2003; 2003WO-US037054.  
 XX 22-NOV-2002; 2002US-0428469P.  
 PR 01-APR-2003; 2003US-0459441P.  
 XX (MEDI-) MEDICAL COLLEGE GEORGIA RES INST.  
 PA (GANA/) GANAPATHY V.  
 PA (INOU/) INOUE K.  
 PA (FEIY/) FEI Y.  
 XX Ganapathy V, Inoue K, Fei Y;  
 XX WPI; 2004-460997/43.  
 DR N-PSDB; ADP64792.  
 XX  
 PT New isolated polynucleotide encoding a Na<sup>+</sup>-coupled citrate transporter  
 PT (NaCT) polypeptide, useful as a drug target for the treatment of obesity,  
 PT hyperlipidemia, and hypercholesterolemia.  
 XX Claim 44; SEQ ID NO 6; 186pp; English.  
 XX The invention relates to novel Na<sup>+</sup>-coupled citrate transporter proteins  
 CC and their encoding genes. Inhibitors of transmembrane citrate  
 CC transporters are useful for extending the lifespan, reducing weight,  
 CC preventing weight gain or lowering blood cholesterol, triglyceride, LDL  
 CC or glucose levels in a subject. The NaCT polypeptide is useful as a drug  
 CC target for the treatment of obesity, hyperlipidemia, and  
 CC hypercholesterolemia. This sequence corresponds to the human Na<sup>+</sup>-coupled  
 CC citrate transporter protein.  
 XX  
 SQ Sequence 568 AA;  
 Query Match 30.2%; Score 875; DB 8; Length 568;  
 Best Local Similarity 35.7%; Pred. No. 1.3e-81;  
 Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;  
 QY 27 IIPILITPLIYGFQDQMAEF-KCLMLIVTMALLWITETPIYVITLFLVFCPLGLV 85  
 DB 19 VTPLLLPLVI-----LMPAKFVRCAYIILMAIYVCTEVIPLAVTSLMPVLLPQLD 74  
 QY 86 ASIVCKQYFDTIIVFLGGLIIVAGIEYSNLHTRIALRVIRIVGSGPRRLFVGLMSVSTF 145  
 DB 75 SRQVCVQYQKMDTNLFLGGLIIVAVVERNLHTRIALRVIRIVGSGPRRLFVGLMSVSTF 134  
 QY 146 MGLWISNSAGTAMCPVIVKALVNELDN-----KIPFVYMTQEEEP-VE 188  
 DB 135 LSMWISNTATTAMVPVIVEAILQOMEATSAATEAGLELVDKGAKEPLGSQVIFEGPTLG 194  
 QY 189 EGEPHPHSKITVAFYAGIAYASSIGGLTIGTGNLVFRGIYTERPPTSTVITFANFM 248  
 DB 195 QOEDQERKRLCKAMTLCICVAASIGGTATLTGTGPNVLLGQMNELFPDSDKLVNFASF 254  
 QY 249 FYSIPLMIVNVTVIITAFIITHMGLFRPN-SKTGKIIAEANTNRKLMEDVLQRHIDL 307  
 DB 255 AFAPFNNLV-----MLLPFWLWLFQVYVNRFNFKKSWGGGLSKKNEKALKVLQEEYRKL 310  
 QY 308 PMSCHEIQMAIAFAPMIVLLITRKPGFVPGWSDLI-----NRKVGSAGSLFIVLLIFAL 363  
 DB 311 PLPSFAEINVLICFLLVILWFSRDPGMPGWLTVAVWVEGETKVS DATVAIFVATLLFIV 370  
 QY 364 PTQYTFKYCC-----GKGPFTAQDAIILSWYVLRNIPWGLLFLGCGFALAVASRET 418  
 DB 371 PSQKPFNFPSQTEERKTPFPY---PPLDVKVTKQKVPWGVLLGGGFALAKSEAS 427  
 QY 419 GLNIMISKAMQVLIGLNPVQVSIITFVLANFFSAFNANVVVAVNIVLPILCEMSLALHLP 478  
 DB 428 GLSVWKGQKQEPHVAVPPAAITLILSLVAVFTCTSNVATTLFLFIPFASMSRSLG 487  
 QY 479 LILTPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTTIIGISVALVNTW 538

Db 488 LYIMLEPCTLSASFAPMLPVATPENAIVFTYGHKLVADMKVTGVIMNIIIGVFCVFLAVNTW 547  
QY 539 GLIIFPEKSFPDWA 553  
Db 548 GRAIP-DLDHFPDWA 561

Search completed: June 30, 2005, 08:52:04  
Job time : 82.2935 secs

**THIS PAGE IS BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:36:57 ; Search time 21.608 Seconds  
(without alignments)  
1941.542 Million cell updates/sec

Title: US-10-017-479A-3

Perfect score: 2897

Sequence: 1 MAEPGEQRKFLVLRCCIFHW.....FPESKSPDWAKEIKNQTKI 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/prodata/1/aaa/5A\_COMB.pep:\*

2: /cgn2\_6/prodata/1/aaa/5B\_COMB.pep:\*

3: /cgn2\_6/prodata/1/aaa/6A\_COMB.pep:\*

4: /cgn2\_6/prodata/1/aaa/6B\_COMB.pep:\*

5: /cgn2\_6/prodata/1/aaa/PTUS\_COMB.pep:\*

6: /cgn2\_6/prodata/1/aaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	920	31.8	180	4	US-09-270-767-42669
2	844	29.1	599	4	US-09-949-016-9866
3	793	27.4	368	4	US-09-270-767-42029
4	743	25.6	626	4	US-09-556-916-20
5	742.5	25.6	627	4	US-09-556-916-8
6	741	25.6	626	4	US-09-556-916-14
7	740.5	25.6	627	4	US-09-556-916-2
8	738.5	25.5	627	4	US-09-949-016-6840
9	737	25.4	626	4	US-09-556-916-22
10	736.5	25.4	627	4	US-09-556-916-10
11	735	25.4	626	4	US-09-556-916-16
12	734.5	25.4	627	4	US-09-556-916-4
13	696	24.0	601	4	US-09-949-016-9977
14	661	22.8	132	4	US-09-270-767-57987
15	624	21.5	580	4	US-09-556-916-24
16	623.5	21.5	581	4	US-09-556-916-12
17	622	21.5	580	4	US-09-556-916-18
18	621.5	21.5	581	4	US-09-556-916-6
19	611	21.1	230	4	US-09-270-767-43713
20	582.5	20.1	561	4	US-09-949-016-8761
21	544	18.8	527	4	US-09-602-787A-516
22	527	18.2	194	4	US-09-270-767-59102
23	453.5	15.7	470	4	US-09-543-681A-5952
24	427.5	14.8	169	4	US-09-270-767-57286
25	351	12.1	233	4	US-09-489-847-176
26	327.5	11.3	335	4	US-09-602-787A-518
27	283.5	9.8	524	3	US-09-134-001C-5457

28	244	8.4	421	4	US-09-248-796A-20749	Sequence 20749, A
29	196	6.8	624	4	US-09-543-681A-4343	Sequence 4343, Ap
30	182	6.3	614	4	US-09-489-039A-12605	Sequence 12605, A
31	179	6.2	302	4	US-09-902-540-10445	Sequence 10445, A
32	164	5.7	430	3	US-09-134-001C-2981	Sequence 2981, Ap
33	162.5	5.6	429	4	US-08-311-731A-287	Sequence 287, App
34	157	5.4	437	4	US-09-543-681A-6984	Sequence 6984, Ap
35	151.5	5.2	694	4	US-09-252-991A-22637	Sequence 22637, A
36	146.5	5.1	443	4	US-09-602-787A-532	Sequence 532, App
37	145.5	5.0	547	4	US-09-489-039A-13843	Sequence 13843, A
38	145	5.0	436	4	US-09-583-110-4729	Sequence 4729, Ap
39	145	5.0	440	4	US-09-107-433-5201	Sequence 5201, Ap
40	144.5	5.0	548	4	US-09-902-540-11870	Sequence 11870, A
41	140	4.8	449	4	US-09-328-352-7512	Sequence 7512, Ap
42	139.5	4.8	596	4	US-09-902-540-13547	Sequence 13547, A
43	132	4.6	525	4	US-09-543-681A-5886	Sequence 5886, Ap
44	132	4.6	751	4	US-09-934-868-58	Sequence 588, Appl
45	131.5	4.5	469	4	US-09-134-000C-5065	Sequence 5065, Ap
46	130	4.5	494	4	US-09-543-681A-7033	Sequence 7033, Ap
47	127.5	4.4	504	4	US-09-543-681A-6790	Sequence 6790, Ap
48	123.5	4.3	462	2	US-08-898-976-2	Sequence 2, Appl1
49	123.5	4.3	462	2	US-08-898-976-4	Sequence 4, Appl1
50	123.5	4.3	478	3	US-09-134-001C-4637	Sequence 4637, Ap
51	123	4.2	493	4	US-09-540-236-2120	Sequence 2120, Ap
52	122.5	4.2	463	4	US-09-252-991A-25488	Sequence 25488, A
53	122	4.2	345	4	US-09-902-540-9836	Sequence 9836, Ap
54	121.5	4.2	478	4	US-09-107-532A-6105	Sequence 6105, Ap
55	120.5	4.2	447	3	US-09-134-001C-4523	Sequence 4523, Ap
56	119.5	4.1	482	4	US-09-328-352-7784	Sequence 7784, Ap
57	118.5	4.1	713	4	US-09-543-681A-6094	Sequence 6094, Ap
58	117	4.0	415	4	US-09-583-110-4250	Sequence 4250, Ap
59	117	4.0	638	4	US-09-252-991A-18036	Sequence 18036, A
60	116.5	4.0	441	4	US-09-489-039A-10612	Sequence 10612, A
61	116	4.0	472	4	US-09-543-681A-5049	Sequence 5049, Ap
62	115	4.0	466	4	US-09-489-039A-7401	Sequence 7401, Ap
63	114.5	4.0	363	4	US-09-270-767-42613	Sequence 42613, A
64	114.5	4.0	485	4	US-09-134-000C-6031	Sequence 6031, A
65	113.5	3.9	470	4	US-09-438-185A-209	Sequence 209, App
66	113	3.9	410	3	US-09-134-001C-5213	Sequence 5213, Ap
67	112.5	3.9	446	4	US-09-543-681A-5864	Sequence 5864, Ap
68	112	3.9	470	4	US-09-328-352-8240	Sequence 8240, Ap
69	112	3.9	477	4	US-09-489-039A-10570	Sequence 10570, A
70	112	3.9	483	4	US-09-489-039A-10360	Sequence 10360, A
71	111.5	3.8	322	4	US-09-134-000C-5510	Sequence 5510, Ap
72	110	3.8	1021	4	US-09-489-039A-13589	Sequence 13589, A
73	109.5	3.8	358	4	US-09-107-532A-5676	Sequence 5676, Ap
74	109.5	3.8	460	4	US-09-252-991A-32040	Sequence 32040, A
75	109.5	3.8	487	4	US-09-489-039A-12980	Sequence 12980, A
76	109.5	3.8	567	4	US-09-543-681A-7515	Sequence 7515, Ap
77	109	3.8	167	4	US-09-248-796A-20740	Sequence 20740, A
78	108.5	3.7	482	3	US-09-134-001C-4309	Sequence 4309, Ap
79	108	3.7	394	4	US-09-107-433-5185	Sequence 5185, Ap
80	108	3.7	408	4	US-09-543-681A-4822	Sequence 4822, Ap
81	108	3.7	471	4	US-09-711-164-444	Sequence 444, App
82	108	3.7	471	4	US-09-492-709A-284	Sequence 284, Appl
83	106.5	3.7	541	3	US-09-614-891-12	Sequence 12, Appl
84	106	3.7	465	3	US-09-134-001C-3526	Sequence 3526, Ap
85	105.5	3.6	447	4	US-09-489-039A-11893	Sequence 11893, A
86	105.5	3.6	466	4	US-09-603-208A-236	Sequence 236, App
87	105.5	3.6	484	4	US-09-489-039A-7690	Sequence 7690, Ap
88	105.5	3.6	496	3	US-09-134-001C-3001	Sequence 3001, Ap
89	105	3.6	466	3	US-09-134-001C-3355	Sequence 3355, Ap
90	105	3.6	467	4	US-09-328-352-5465	Sequence 5465, Ap
91	105	3.6	492	4	US-09-134-000C-4808	Sequence 4808, Ap
92	104.5	3.6	408	4	US-09-328-352-5768	Sequence 5768, Ap
93	104.5	3.6	473	4	US-09-710-279-60	Sequence 60, Appl
94	104.5	3.6	555	4	US-09-543-681A-4582	Sequence 4582, Ap
95	104.5	3.6	797	4	US-09-134-000C-4997	Sequence 4997, Ap
96	104	3.6	327	4	US-09-949-016-7336	Sequence 7336, Ap
97	104	3.6	331	4	US-09-134-000C-6008	Sequence 6008, Ap
98	104	3.6	499	4	US-09-328-352-5089	Sequence 5089, Ap
99	103.5	3.6	633	4	US-09-248-796A-20407	Sequence 20407, A
100	103	3.6	471	4	US-09-543-681A-7497	Sequence 7497, Ap

## ALIGNMENTS

## RESULT 1

US-09-270-767-42669  
; Sequence 42669, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 42669  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: *Drosophila melanogaster*  
US-09-270-767-42669

Query Match 31.8%; Score 920; DB 4; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.3e-84;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 383 AIDAILSWYVLRNIPWGLFLGCGFALAVASRETGLNIMISKAMQVLGLNIVVQSI 442  
Db 1 AIDAILSWYVLRNIPWGLFLGCGFALAVASRETGLNIMISKAMQVLGLNIVVQSI 60  
  
Qy 443 TFVLNFFSANVNVANIVLPILCENSLALELHPLILTPACLGISMVYFLPVSTPPN 502  
Db 61 TFVLNFFSANVNVANIVLPILCENSLALELHPLILTPACLGISMVYFLPVSTPPN 120  
  
Qy 503 AIVTQYAHIKTKYPACCGIVPTIIGISVALVNTNTWGLIIFPESKSPDWAKEIKNOTKI 562  
Db 121 AIVTQYAHIKTKYPACCGIVPTIIGISVALVNTNTWGLIIFPESKSPDWAKEIKNOTKI 180

## RESULT 2

US-09-949-016-9866  
; Sequence 9866, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9866  
; LENGTH: 599  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9866

Query Match 29.1%; Score 844; DB 4; Length 599;  
Best Local Similarity 33.9%; Pred. No. 3.5e-76;  
Matches 200; Conservative 111; Mismatches 199; Indels 80; Gaps 14;  
  
Qy 20 WRGKASII---IPILITPILYIGFTDWAEPKCLWLIVTALLWITETLPYVTALEPLV 76  
Db 16 WAYRSYLIVFVFPILLPLPIL---VPSKEAYCAYAILMALFWCTEALPLAVTALPFLI 72

Qy 77 FCPLLGLVNASIVCKQYFTDTIIVFLGGLIVLVALGIEYSNLHTRIALRVIRIVGSGPRRLF 136  
Db 73 LFPFMGIVDASEVAVEYLDXSNLLFFGGLLVAIVEHNLHKEIALRVLLIVGVRAPLI 132  
  
Qy 137 VGLMSVSTFMGLWISNSAGTAMMCPIVKALVNELDNKNIPVYMTQEEBPVEG----- 190  
Db 133 LGFMLVTAFLSMWISNTATSAMVPIAHAVLDQ-----HSSQASSNVESGNNPTF 184  
  
Qy 191 --EPHPHSK-----ITVAFYAG-----IYASSIGIGLGL 218  
Db 185 ELQEPSQKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCMSLCVCSYASIGGLATL 244  
  
Qy 219 IGTGTLNLFERGIYTERFPTSTVEITPANFMFYSIPLMWIVNVTLVIAFL---ITHMGL- 274  
Db 245 TGTAPNLVLQGGINSILFPQNGVNVNFAFSFAFTWVI-----LLLLLWMLQLFLGFN 300  
  
Qy 275 FRNSKTGKIIABANTNRKLMEDVLRQRHIDGPMSCHEIQMAIAFAMVLVLTITKPGF 334  
Db 301 FRKNFGIGEKMQE---QQAAYCVIQTHERLLGPMTFAEKATISILEFVILVLLWFTREPGF 357  
  
Qy 335 VPGWSDLI-----NRKVVGSAGLSFIVLLIPALPTQYTFYKCCG-----KGPFTAQAI 384  
Db 358 FLGWGNLAPNNAKGSWSDGVAIFIGIIMFIIPSKFP-----GLTQDPENPGKLLKAP 411  
  
Qy 385 DAILSWYVLRNIPWGLFLGCGFALAVASRETGLNIMISKAMQVLGLNIVVQSI 444  
Db 412 LGLDWTNQRKMPNIVLLGGYALAKGSGSLSEWLNKLTPLQSVAPAPAIILS 471  
  
Qy 445 VLANFFSANVNVANIVLPILCENSLALELHPLILTPACLGISMVYFLPVSTPPN 504  
Db 472 LLVATFECTSNVATTTIFLPILASMAQAICLHPLYVMLPCTLTATSLAFMLPVATPPN 531  
  
Qy 505 VITQYAHIKTKYPACCGIVPTIIGISVALVNTNTWGLIIFPESKSPDWAKEIKNOTKI 554  
Db 532 VPSFGDLKVDMARAGFLNIIIGVLIILAINSWGIPLP-SLHSFFSWAQ 580

## RESULT 3

US-09-270-767-42029  
; Sequence 42029, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 42029  
; LENGTH: 368  
; TYPE: PRT  
; ORGANISM: *Drosophila melanogaster*  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-42029

Query Match 27.4%; Score 793; DB 4; Length 368;  
Best Local Similarity 43.6%; Pred. No. 2.4e-71;  
Matches 154; Conservative 71; Mismatches 112; Indels 16; Gaps 6;  
  
Qy 3 EPGEQRKFLVGRCCIF---HWRGKASIIIPILITPILYIGFTDWAEPKCLWLIVTALL 59  
Db 17 EIGEQQPPV-KCSNFFANHWKGLVFLVPLLCPLVMLL--NEGAEFRCWMLLVNAIF 72  
  
Qy 60 WITETLPYVTALEPLVFCPLLGLNIVCKQYFTDTIIVFLGGLIVLVALGIEYSNLHTR 119  
Db 73 WVTALPLVYVTSIMPIVAFPIIMGIMSSDQTCRLYFKDTLVFMFGGINVALAVEYCNLHNR 132  
  
Qy 120 IALRVIRIVGSGPRRLFVGLMSVSTFMGLWISNSAGTAMMCPIVKALVNELDNKNIPVY 179  
Db 133 LAURVIXXXXXXXXXXXXXXXXXXXXXXMMWISNACTMMCPILQAVLEELQAGVCKIN 192







```

Db      605 VSLF-HLDTYPAPAWAR-VSNIT 623

RESULT 10
US-09-556-916-10
; Sequence 10, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-10

Query Match      25.4%; Score 736.5; DB 4; Length 627;
Best Local Similarity 28.0%; Pred. No. 2.6e-65;
Matches 174; Conservative 124; Mismatches 223; Indels 101; Gaps 14;

QY 26 IIPILITPLIYGFQTDMAEFKCLMLIVTMALLMITETPLIYVVTALFPLVFCPLLLGLVN 85
Db 17 VCVPLLLPLPVL---HPSSSEACAYLVITAVYVWSEAVPLGAAALVPALFPPFGVL 73
QY 86 ASIVCKQYFTDTIVFVGLGILVALGIEYSNLHRIALRVIRIVGSGPRRLFVGLMSVSTF 145
Db 74 SNEVAEYFKNTLLLVGVICVAAAEKVNHLKRIALRMVLMAGAKPGMLLLCFMCCCTL 133
QY 146 MGLWISNSAGTAMMCPVKALVNEL-----DTNKIPFVYMTQEEPV----- 187
Db 134 LSMWLSNTSTTAMVMPIVEAQLVQELVSAEDEQLVAGNSNTEAEAPISLDVKNQSLSLELI 193
QY 188 -----ELEGPPHPSK----- 197
Db 194 FVNEDRSNADLTLMHNENLVPSITNPIKTANQHQKQHPSEKQPVLTPSPRKQLN 253
QY 198 -----ITVAFVAGIAYASSIGGLTGLTGTGNLVFRGIYTERFPTSTVEITFANF 246
Db 254 NRKYRSHDDQMI CKLSLSISYSATIGGLTIIIGTSTSLIFLEHFNNOYPAAEV-VNFGT 312
QY 247 FMFYSIPLMIVNVTVIIAFLITHMGLFRPNSK-TGKIIEAENTNR-KLMEDVLRQRIH 304
Db 313 WFLFSPFISLI-----MLVVSFWFMHMLFLGCNFKETCSLSKKKTKREQLSEKRIQEEYE 368
QY 305 DLGPMSCHEIOMAIAFAMFVLLITRKPGVPWGSDDLNRKVVGSASGLS-FIVLLIFAL 363
Db 369 KLGDISPEVMVTGFFFLMTVLTFTREPFGVPGWDSFFEKKGYRTDATSVFLGFLLELI 428
QY 364 PTQYTFPKYCCGK---GPFFTAQAI---DAILSWEVLRNIPWGLLFLGGGFALAVASRET 418
Db 429 PAK-----KPCFGKKNNGENQEHSLGTSIITWKDFQKTMPEWIVLVGGGYALASGSKS 484
QY 419 GLNIMISKAMQVLIGLPNI VQSGITFVLNFFSAFNANVVVAVIVLPILCEMSLALHLHP 478
Db 485 GLSTWIGNQMLSSLSLPPWAVTLLACILVSIVTFVSNPATITIFLPILCSLSETOHNP 544
QY 479 LIILTPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGVPTIIGTSVALVNTNW 538
Db 545 LYTLIPVTMCISFAVMLPVGNPNNAIVFSYGHQCIQKDMVYAGLVGNVIGLIVVMVAINTW 604
QY 539 GLIIFPESKFPDPAWEIKNQ 560
Db 605 GVSLF-HLDTYPAPAWAR-VSNIT 624

RESULT 11
US-09-556-916-16
; Sequence 16, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-16

Query Match      25.4%; Score 735; DB 4; Length 626;
Best Local Similarity 28.0%; Pred. No. 3.6e-65;
Matches 174; Conservative 123; Mismatches 224; Indels 100; Gaps 14;

QY 26 IIPILITPLIYGFQTDMAEFKCLMLIVTMALLMITETPLIYVVTALFPLVFCPLLLGLVN 85
Db 17 VCVPLLLPLPVL---HPSSSEACAYLVITAVYVWSEAVPLGAAALVPALFPPFGVL 73
QY 86 ASIVCKQYFTDTIVFVGLGILVALGIEYSNLHRIALRVIRIVGSGPRRLFVGLMSVSTF 145
Db 74 SNEVAEYFKNTLLLVGVICVAAAEKVNHLKRIALRMVLMAGAKPGMLLLCFMCCCTL 133
QY 146 MGLWISNSAGTAMMCPVKALVNEL-----DTNKIPFVYMTQEEPV----- 187
Db 134 LSMWLSNTSTTAMVMPIVEAQLVQELVSAEDEQLVAGNSNTEAEAPISLDVKNQSLSLELI 193
QY 188 -----ELEGPPHPSK----- 197
Db 194 FVNEDRSNADLTLMHNENLVPSITNPIKTANQHQKQHPSEKQPVLTPSPRKQLN 253
QY 198 -----ITVAFVAGIAYASSIGGLTGLTGTGNLVFRGIYTERFPTSTVEITFANF 247
Db 254 NRKYRSHDDQMI CKLSLSISYSATIGGLTIIIGTSTSLIFLEHFNNOYPAAEV-VNFGTW 312
QY 248 FMFYSIPLMIVNVTVIIAFLITHMGLFRPNSK-TGKIIEAENTNR-KLMEDVLRQRIH 305
Db 313 WFLFSPFISLI-----MLVVSFWFMHMLFLGCNFKETCSLSKKKTKREQLSEKRIQEEYE 368
QY 306 LQPMSCHEIOMAIAFAMFVLLITRKPGVPWGSDDLNRKVVGSASGLS-FIVLLIPALP 364
Db 369 LQDISPEVMVTGFFFLMTVLTFTREPFGVPGWDSFFEKKGYRTDATSVFLGFLLELI 428
QY 365 TQYTFPKYCCGK---GPFFTAQAI---DAILSWEVLRNIPWGLLFLGGGFALAVASRET 419
Db 429 AK-----KPCFGKKNNGENQEHSLGTEPIITWKDFQKTMPEWIVLVGGGYALASGSKSG 484
QY 420 LNMISKAMQVLIGLPNI VQSGITFVLNFFSAFNANVVVAVIVLPILCEMSLALHLHP 479
Db 485 LSTWIGNQMLSSLSLPPWAVTLLACILVSIVTFVSNPATITIFLPILCSLSETOHNP 544
QY 480 IITLPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGVPTIIGTSVALVNTNW 539
Db 545 YTLIPVTMCISFAVMLPVGNPNNAIVFSYGHQCIQKDMVYAGLVGNVIGLIVVMVAINTW 604
QY 540 LIIFPESKFPDPAWEIKNQ 560
Db 605 VSLF-HLDTYPAPAWAR-VSNIT 623
```



```
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57987
; LENGTH: 132
; TYPE: PRN
; ORGANISM: Drosophila melanogaster
US-09-270-767-57987

Query Match      22.8%; Score 661; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-58;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 383 AIDALLSWEVLRNIPWGLLFLGGGFALAVASRETGLNIMISKAMQVLGLPNIVVQSI 442
Db 1 AIDALLSWEVLRNIPWGLLFLGGGFALAVASRETGLNIMISKAMQVLGLPNIVVQSI 60

QY 443 TFVLNPFSAFNANVVANIVLPILCEMSLALHPLILTLPAICLGISMVYFLPVSTPPN 502
Db 61 TFVLNPFSAFNANVVANIVLPILCEMSLALHPLILTLPAICLGISMVYFLPVSTPPN 120

QY 503 AIVTOYAHIKTK 514
Db 121 AIVTOYAHIKTK 132
```

RESULT 15

```
US-09-556-916-24
; Sequence 24, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 580
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-556-916-24
```

```
Query Match      21.5%; Score 624; DB 4; Length 580;
Best Local Similarity 25.4%; Pred. No. 5.1e-54;
Matches 158; Conservative 121; Mismatches 196; Indels 146; Gaps 16;

QY 26 IIPILITLPIILYGFQDIAEFKCLMLIVTMALLWITETLPYVVTALFPLVFCPLLGLVN 85
Db 17 VCVPLLLPLPVL---HPSSEASCAVYLIIVTAVYVWSEAVPLGNAALVPAPFLYFFGVLR 73

QY 86 ASIVCKQYFDTIIVFLGGLIVAGIEYSNLHTRIALRVIRIVGGSRRRLFVGLMSVSTF 145
Db 74 SNEVAAEYFNKNTLLLVGVICVAAAEKLNHLKRIALRMVLMAGAKPGMLLLCFMCCTIL 133

QY 146 MGLWISNSAGTAMMCPVIALVNL-----DTNKIFPVYMTQEEBPV----- 187
Db 134 LSMWLSNTSTTAMVMPVIEAVLQELVSAEDEQLVAGNSNTEEAEPISLDVKNQPSLELI 193

QY 188 -----EEGEPHPPSK----- 197
Db 194 FVNEESNADLTLMNENLNGVPSITNPIKTANQHQKQHPQKQVLTPTSPRKQKLN 253

QY 198 -----ITVAFYAGIAYASSIGGLTGLTGTLNLFVFRGIYTERPFTSTVEITFANF 247
```

```
Db 254 RKVRGHHQDMICKCLSLSSISYSATIGGLTIIGTSTSLIFLHFNNQYPAAEV-VNFGTW 312
QY 248 MEYSIPLMVIIVNVTVIIAFLITHMGLRPNPSK-TGKIIABEANTNR-KLMEDVLRORHID 305
Db 313 FLFSFPISLI-----MLVVSFWFMHMLFLGCNFKETCSLSKKKKTKREQLSEKRIQBEYEK 368
QY 306 LGPMSCHETQMAIAFAFMIVLITRKPGFVPGWSDLINRKVVVGSASGLS-FIVLILFALP 364
Db 369 LGDISYPMVGTGFFFLMTVLMFTREPFGVPCWDSFFEKKGYRTDATVSVFLGLFLIP 428
QY 365 TQYTFKYCCGK---GPFTAQAI--DAILSWBYVLRNIPWGLLFLGGGFALAVASRETG 419
Db 429 AK---KPCFGKKNDCGENQESLGTESIITWKDFOKTMEPWEIVILVGGGYALASGSKSG 484
QY 420 LNMISKAMQVLGLPNIVVQSIITFVLNPFSAFNANVVANIVLPILCEMSLALHPL 479
Db 485 LSTWIGNQMLSSLPFWAVTLLACILSVITFEVSNPATITIFLPLCSL----- 535
QY 480 ILTLPAICLGISMVYFLPVSTPPNNAIVTOYAHIKTKYFACCGIVPTIIGISVALVNTWTG 539
Db 536 ---VKAGLGVN-----VIGLVIVMVAINTWG 558
QY 540 LIIFPESKSFPPDWAKEIKNOT 560
Db 559 VSLF-HLDTYPAPAWAR-VSNIT 577

Search completed: June 30, 2005, 09:02:13
Job time : 23.608 secs
```



```
85 251 8.7 162 11 US-09-833-245-1945 Sequence 1945, Ap
86 250.5 8.6 236 16 US-10-767-701-43427 Sequence 43427, A
87 224.5 7.7 174 16 US-10-425-115-335649 Sequence 335649,
88 223.5 7.7 226 15 US-10-296-115-980 Sequence 980, App
89 220.5 7.6 300 16 US-10-767-701-46317 Sequence 46317, A
90 214.5 7.4 259 16 US-10-739-930-9458 Sequence 9458, Ap
91 207 7.1 78 9 US-09-864-761-34252 Sequence 34252, A
92 207 7.1 78 14 US-10-029-386-28038 Sequence 28038, A
93 207 7.1 163 15 US-10-424-599-184881 Sequence 184881,
94 206.5 7.1 234 15 US-10-335-977-5397 Sequence 5397, Ap
95 185 6.4 457 15 US-10-369-493-4366 Sequence 4266, Ap
96 180 6.2 429 16 US-10-482-706-191 Sequence 191, App
97 178 6.1 135 15 US-10-276-774-1592 Sequence 1592, Ap
98 169.5 5.9 494 15 US-10-335-977-5882 Sequence 5882, Ap
99 169 5.8 164 15 US-10-424-599-205152 Sequence 205152,
100 165.5 5.7 482 15 US-10-335-977-5881 Sequence 5881, Ap
```

ALIGNMENTS

```
RESULT 1
US-10-167-994-10
; Sequence 10, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 562
; TYPE: PRP
; ORGANISM: Drosophila melanogaster
US-10-167-994-10
```

```
Query Match 95.7%; Score 2772; DB 14; Length 562;
Best Local Similarity 92.0%; Pred. No. 5.4e-250;
Matches 519; Conservative 39; Mismatches 2; Indels 4; Gaps 2;

Qy 1 MAEPGEQRKFFVLGRCCIFHWRGKASIIPLITLPIIYGFQTDMAEFKCLWLIVTMALLW 60
Db 1 MAEPGEQRKFFVLGRCCIFHWRGKASIIPLITLPIIYGFQTDMAEFKCLWLIVTMALLW 60

Qy 61 ITETLPIYVTALFPLVFCPLGLVNASIVCKQYFTDTTIVFVLGLI VALGIEYSNLHTRI 120
Db 61 L TETLPIYVTALFPLVFCPLGLVNASIVCKQYFTDTTIVFVLGLI VALGIEYSNLHTRI 120

Qy 121 ALRVIRIVGGSPRRLLFVGLMSVTFMGLWINSAGTAMMCPIVKALVNELDTNKLFPVYM 180
Db 121 ALRVIRIVGGSPRRLLFVGLMSVTFMGLWINSAGTAMMCPIVKALVNELDTNKLFPVYM 180

Qy 181 TOEEEPVEEGPPHPSKITTAFYAGIAYASSIGGLGTGLTGTLNLPFRGIYTERPPTSTV 240
Db 181 TOEEEPVEEGPPHPSKITTAFYAGIAYASSIGGLGTGLTGTLNLPFRGIYTERPPTTV 240

Qy 241 EITFANFMFYSIPLMVIYNVTLLVIAFLITHMGLFRPNSTKTKIAEANTNRKLMEDVLR 300
Db 241 EITFANFMFYSIPLMVIYNVTLLVIAFLITHM -FRPNSTKTKIAEANTNRKLMEDVLR 298

Qy 301 QRHDLGPMSCHEIQMAIAFAPMIVLLTRKPGFVPGHSDLINRKVGSASGLSFIVLLI 360
Db 299 QRHDLGPMSCHEIQMALAFAPMIVLLTRKPGFVPGHSDLINRKVGSASGLSFIVLL 358
```

```
Qy 361 FALPTQYTFYFKYCC--GKGPFYTAQADAILSWYVLRNI PWGLLFLGGGFALAVASRET 418
Db 359 FALPTQYTFYFKYCCGKGGPFYTAQADAILSWYVLRNI PWGLLFLGGGFALAVASRET 418

Qy 419 GUNIMSKAMQVLIGLPMIVQSIITFVLANPFSANFANVAVNI VLPILCENSLELHP 478
Db 419 GUNIMSKAMQVLIGLPMIVQSIITFVLANPFSANFANVAVNI VLPILCENSLELHP 478

Qy 479 LILTPACILGISMVYFLVPSTPPNAIVTQYAHKTKYFACCGIVPTIIGISVALVNTNTW 538
Db 479 LILTPACILGISMVYFLVPSTPPNAIVTQYAHKTKYFACCGIVPTIIGISVALVNTNTW 538

Qy 539 GLIIPESKSFDPDAKEIKNQTKI 562
Db 539 GLIIPESKSFDPDAKEIKNQTKL 562

RESULT 2
US-10-167-994-2
; Sequence 2, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 572
; TYPE: PRP
; ORGANISM: Drosophila melanogaster
US-10-167-994-2
```

```
Query Match 47.7%; Score 1382.5; DB 14; Length 572;
Best Local Similarity 46.5%; Pred. No. 4.2e-120;
Matches 262; Conservative 120; Mismatches 164; Indels 17; Gaps 7;

Qy 3 EPGEQRKFFVLGRCCIF---HWRGKASIIPLITLPIIYGFQTDMAEFKCLWLIVTMALL 59
Db 4 EIGEQQPPV-KCSNFFANHWKGLVFLVPLCLLPVLL---NEGAEFRCMYLLVMAIF 59

Qy 60 WITETLPIYVTALFPLVFCPLGLVNASIVCKQYFTDTTIVFVLGLI VALGIEYSNLHTR 119
Db 60 WYTEALPLVYTSIMPIVAFPIMGINSDDQTCRLYFKDITLVFMFGGIMVALAVEYCNLHTR 119

Qy 120 IALRVIRIVGGSPRRLLFVGLMSVTFMGLWINSAGTAMMCPIVKALVNELDTNKLFPVY 179
Db 120 LALRVIRIVGGSPRRLLHFGIMVTFMGLWINSAGTAMMCPIIQAVLELOAQGVCKIN 179

Qy 180 MTQEEPPV-----EAGEPPHPSKITVAFYAGIAYASSIGGLGTGLTGTLNLPFRGIYTE 233
Db 180 HEQYQIVGQGNKKNEDPEPYTKITLCYVLGIAYASSIGGCGTIIGTATNLAFKGIYE 239

Qy 234 RPTSTVEITFANFMFYSIPLMVIYNVTLLVIAFLITH-MGLFRPNSTKTKIAEANTNR 292
Db 240 RFKNSTEQMDPFTFMFYSVPSMLV--YTLTTFVFLQWHFMGLWRPKSKEAQVQREGA 297

Qy 293 KLMEVLRQRHDLGPMSCHEIQMAIAFAPMIVLLTRKPGFVPGHSDLINRKVGSASG 352
Db 298 DVAKKVIDQRYKOLGPMSCHEIQMILIFIVVMVYTRKPGIFLWADLNSKDIRNSMP 357

Qy 353 LSFIVLLI FALPTQYTFYFKYCCGK--GPFTAQADAILSWYVLRNI PWGLLFLGGGFAL 411
Db 358 TIFVVMCFMLPANYAFRLYCTRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFAL 417
```

QY 412 AVASRETLNIMISKAMQVLIGLNPVQSIPTVLNFFSAFNANVAVNIVLPILCEMS 471  
Db 418 AEGSKQSMAGLIGNALIGLKLNPNSVLLWLVAVFLTAFSSNVAIANIIPVLAEMS 477  
QY 472 LALEHPLILTPACLGISVMVFLPVSTPPNNAIVTOYAHKTKYFACCGIVPTIIGISVA 531  
Db 478 LAIEHPLYLILPAGLACSMAFHLPVSTPPNALVAGYANIRYKDMAIAGIGPTIITL 537

QY 532 LVNTNTWGLIIPESKSPFDWAK 554  
Db 538 FVFCQTWGLVVPNLNSFPPEWAQ 560

## RESULT 3

US-10-017-479-2  
; Sequence 2, Application US/10017479  
; Publication No. US20030104399A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Connecticut  
; APPLICANT: Helfand, Stephan L  
; APPLICANT: Reenan, Robert A  
; APPLICANT: Rogina, Blanka  
; TITLE OF INVENTION: Polynucleotides Encoding Cellular Transporters and Methods of Use  
; FILE REFERENCE: UCT-0020  
; CURRENT APPLICATION NUMBER: US/10/017,479  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: 60/255,013  
; PRIOR FILING DATE: 2000-12-12  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 2  
; LENGTH: 572  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-017-479-2

Query Match 47.7%; Score 1382.5; DB 14; Length 572;  
Best Local Similarity 46.5%; Pred. No. 4.2e-120;  
Matches 262; Conservative 120; Mismatches 164; Indels 17; Gaps 7;

QY 3 EPGEQKFLVGRCCIF---HWRGKASIIPLITLPIIYGFOTDMAEFKCLWLVTMALL 59  
Db 4 EIGEOPQPVV-KCSNFFANHWKGLVFLVPLCLPVMLL---NEGAEFRMCLLVMAIF 59  
QY 60 WITETLPIYVLTALPFLVFCPLGLGNASIVCKQYFTDTIVVFLGGLI VALGIEYENLHTR 119  
Db 60 WYTEALPLYVTSMIPVAFPIGIMSSDOTCRLFKDTLVMPMGIMVALAVEYCNLHXR 119  
QY 120 IALRVIRIVGSPRLFLVGLMSVTFMGLWISNSAGTAMMCPV KVALVNELDTNKIFPVY 179  
Db 120 LALRVIOIVGCSPRRLHFLIMVTFMGLWISNSAGTAMMCPV KVALVNELDTNKIFPVY 179  
QY 180 MTQEEEPV-----EGBEPHPKTI VAFYAGIAYASSIGLGTIGTNLVPRGIYTE 233  
Db 180 HEPOYQIVGGNKNNEDEPPYTKITLCYLGIAIYASSLGGCGTIGTATNLTFKGIYEA 239  
QY 234 RPPTSTVEITFANFMFYSIPLMVIVNVLTVIIAFLITH-MGLFRPNSTKGIKIAEANTNR 292  
Db 240 RFKNSTQMDPTFMFYSVPSMLV--YTLTFVFLQWFMGLWRPKSKAEQVORREGA 297  
QY 293 KLMDVLQRHIDLGPMSCHEIQMAIAFAMVILLITRKPGFVPGWSDLINRKVVGSAG 352  
Db 298 DVAKKVIDQRYKDLGPMSCHEIQMILFIMVMYFTRKPGIFLGWADLLNSKDIRNSMP 357  
QY 353 LSFVILLIFALPTQYTFKCYCGK-GPETAQIDAISWEYVLNIPMGLLFLGGG 411  
Db 358 TIFVVMCFMPLPANYAFRLYCTRGGVPVPTGPTSLITWKFIQTKVPMGLVFLGGG 417  
QY 412 AVASRETLNIMISKAMQVLIGLNPVQSIPTVLNFFSAFNANVAVNIVLPILCEMS 471  
Db 418 AEGSKQSMAGLIGNALIGLKLNPNSVLLWLVAVFLTAFSSNVAIANIIPVLAEMS 477  
QY 472 LALEHPLILTPACLGISVMVFLPVSTPPNNAIVTOYAHKTKYFACCGIVPTIIGISVA 531  
Db 478 LAIEHPLYLILPAGLACSMAFHLPVSTPPNALVAGYANIRYKDMAIAGIGPTIITL 537

Db 478 LAIEHPLYLILPAGLACSMAFHLPVSTPPNALVAGYANIRYKDMAIAGIGPTIITL 537  
QY 532 LVNTNTWGLIIPESKSPFDWAK 554  
Db 538 FVFCQTWGLVVPNLNSFPPEWAQ 560

## RESULT 4

US-10-718-359-2  
; Sequence 2, Application US/10718359  
; Publication No. US20050095240A1  
; GENERAL INFORMATION:  
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.  
; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION  
; FILE REFERENCE: 275,0008,0101  
; CURRENT APPLICATION NUMBER: US/10/718,359  
; CURRENT FILING DATE: 2003-11-20  
; PRIOR APPLICATION NUMBER: 60/428,469  
; PRIOR FILING DATE: 2002-11-22  
; PRIOR APPLICATION NUMBER: 60/459,441  
; PRIOR FILING DATE: 2003-04-01  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 2  
; LENGTH: 572  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-718-359-2

Query Match 47.7%; Score 1382.5; DB 17; Length 572;  
Best Local Similarity 46.5%; Pred. No. 4.2e-120;  
Matches 262; Conservative 120; Mismatches 164; Indels 17; Gaps 7;

QY 3 EPGEQKFLVGRCCIF---HWRGKASIIPLITLPIIYGFOTDMAEFKCLWLVTMALL 59  
Db 4 EIGEOPQPVV-KCSNFFANHWKGLVFLVPLCLPVMLL---NEGAEFRMCLLVMAIF 59  
QY 60 WITETLPIYVLTALPFLVFCPLGLGNASIVCKQYFTDTIVVFLGGLI VALGIEYENLHTR 119  
Db 60 WYTEALPLYVTSMIPVAFPIGIMSSDOTCRLFKDTLVMPMGIMVALAVEYCNLHXR 119  
QY 120 IALRVIRIVGSPRLFLVGLMSVTFMGLWISNSAGTAMMCPV KVALVNELDTNKIFPVY 179  
Db 120 LALRVIOIVGCSPRRLHFLIMVTFMGLWISNSAGTAMMCPV KVALVNELDTNKIFPVY 179  
QY 180 MTQEEEPV-----EGBEPHPKTI VAFYAGIAYASSIGLGTIGTNLVPRGIYTE 233  
Db 180 HEPOYQIVGGNKNNEDEPPYTKITLCYLGIAIYASSLGGCGTIGTATNLTFKGIYEA 239  
QY 234 RPPTSTVEITFANFMFYSIPLMVIVNVLTVIIAFLITH-MGLFRPNSTKGIKIAEANTNR 292  
Db 240 RFKNSTQMDPTFMFYSVPSMLV--YTLTFVFLQWFMGLWRPKSKAEQVORREGA 297  
QY 293 KLMDVLQRHIDLGPMSCHEIQMAIAFAMVILLITRKPGFVPGWSDLINRKVVGSAG 352  
Db 298 DVAKKVIDQRYKDLGPMSCHEIQMILFIMVMYFTRKPGIFLGWADLLNSKDIRNSMP 357  
QY 353 LSFVILLIFALPTQYTFKCYCGK-GPETAQIDAISWEYVLNIPMGLLFLGGG 411  
Db 358 TIFVVMCFMPLPANYAFRLYCTRGGVPVPTGPTSLITWKFIQTKVPMGLVFLGGG 417  
QY 412 AVASRETLNIMISKAMQVLIGLNPVQSIPTVLNFFSAFNANVAVNIVLPILCEMS 471  
Db 418 AEGSKQSMAGLIGNALIGLKLNPNSVLLWLVAVFLTAFSSNVAIANIIPVLAEMS 477  
QY 472 LALEHPLILTPACLGISVMVFLPVSTPPNNAIVTOYAHKTKYFACCGIVPTIIGISVA 531  
Db 478 LAIEHPLYLILPAGLACSMAFHLPVSTPPNALVAGYANIRYKDMAIAGIGPTIITL 537  
QY 532 LVNTNTWGLIIPESKSPFDWAK 554  
Db 538 FVFCQTWGLVVPNLNSFPPEWAQ 560

```
RESULT 5
US-10-092-900A-44
; Sequence 44, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Caeman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE OF INVENTION: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/992,900A
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 44
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-900A-44
Query Match 30.5%; Score 885; DB 15; Length 568;
Best Local Similarity 35.6%; Pred. No. 1.3e-73;
Matches 197; Conservative 101; Mismatches 219; Indels 36; Gaps 9;
```

```
QY 27 IIPILITLIIYGFQTDMAEFKCLMLIVTMALLWITETLPIVVTALFPLVFCPLLGLVNA 86
DB 19 VTPLLLPLVI-----LMPAKVSCAVIILMAYTCVTEVPLAVTSLMPVLLPLFQLILDS 74
QY 87 STVCKQYFTDTIVVFLGGLIIVAGLIEYSLNHPRIALRIVIRIVGGSPRRRLFVGLMSVSTM 146
DB 75 RQVCVQYMKDTNMLFLGGLIIVAVAVERNLHKRIALRTLTLVWGAKPARLMLGFMGVTTALL 134
QY 147 GLWISNSAGTAMCMCIIVKALVNELDN-----KIPVYVMTQEEER-VSE 189
DB 135 SMWISNTATTAMVPIVEAILQOOMEATSATBAGLEVDKGAKEPLPGSQVIFEGTTLQ 194
QY 190 GEPPHPSKITVAFYAGIAYASSIGGLTGLTGTNLVFRGIYTERPPTSTVEITTFANFME 249
DB 195 QEDQERKRLCKAMTLCICVNASIGGTATLTGTGNVLLGQNNELFPDPSKDLVNFASWFA 254
QY 250 YSIPLMVIVNVTLVIAFL-ITHMGLFRPNSKTGKIIABANTNRKJMEDVLRQRHIDLGP 308
DB 255 FAPPNNLV---MLLFAWLWLQFVYMFSSFKKSWGCGLESKKEKAALKVLOEYRKLGP 310
QY 309 MSCHETOMAIAPAFMVLVILITRKPQFVPCGWSDLI----NRKVVGSAAGLSFVILLIFALP 364
DB 311 LSAEINVLICPFLVILWFSRDPGFMGWLTVAVWVEGETKTVSDATVAIFVATLFIIVP 370
QY 365 TOVTEPKYCC---GKGPTAQDAIDAILSWEVYLRNIPMGLLFLGGLGFALAVASRETGL 420
DB 371 SQPKENFRSQTEEGKSPVLI-APPPLLDKVKYQEKVPMGIVLLGGLGFALAKGSEASGL 429
QY 421 NIMISKAMQVLIGLPNIVVOSTITFVLNFFSAPFANVNVANIVLPILCMSLALHPLI 480
DB 430 SVMMGQMEPLHAPPAITLILSLVAVFETCTSNVATTTFLPPIFASMSRSIGLNPLY 489
QY 481 LTLPACLGISMVYFLPVSPTPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTNTWGL 540
DB 490 IMLPCTLSASFAPMLPVATPPNAIVFTYGLKVDKVMKGVINNIIOVFCVFLAVNTWGR 549
QY 541 IIPPEKSPDPDA 553
DB 550 AIF-DLDHFPDWA 561
RESULT 6
US-10-403-161-70
; Sequence 70, Application US/10403161
; Publication No. US20040043930A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-573C
; CURRENT APPLICATION NUMBER: US/10/403,161
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/370349
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384543
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/370969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/403748
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/372019
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/374379
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 09/779679
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 10/055877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262892
; PRIOR FILING DATE: 2001-01-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 70
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-161-70

Query Match      30.5%; Score 885; DB 15; Length 568;
Best Local Similarity 35.6%; Pred. No. 1.3e-73;
Matches 197; Conservative 101; Mismatches 219; Indels 36; Gaps 9;

QY 27 IIPLTILPILYGFQTDMAEFKCLMLIVTMALLWITETLPIYVYVTFALPFLVFCPLGLVNA 86
Db 19 VTPLLLPLVI-----LMPAKFVRCAYVILMAIYVCTEVIPLAVTSLMPVLLPFLQILD 74
QY 87 SVVCKQYFDTTIVFLGGLIVALGIEYSLHTRIALRVIRVGGSPRRFLVGLMSVTFM 146
Db 75 ROVCVQYKMDTNMLFLGGLIIVAVAVERNLHRIALRTLWVGAKPARMLMGFGVTALL 134
QY 147 GLWISNSAGTAMMCPIVKALVNELDTN-----KIPPVYMTQEEEP-VVE 189
Db 135 SMWISNTATTAMVPIVEAILQOMEATSAATEAGLELDVKGKAKELPGSQVIFEGPTLQ 194
QY 190 GEPHPHSKITVAFYAGIAYASSIGGLTGLTGTLNLFVFRGIYTERPPTSTVETITANFM 249
Db 195 QEDQERKRLCKAMTLCICYAASIGGTATLTGTGNVVLGOMNELFPDSKOLVNEASWFA 254
QY 250 YSIPLMVINVTAVIAEL-ITHMGLFRNSKTGIIABANTNRKLMEDVLRQRIHIDLP 308
Db 255 FAFPNMLV-----MLLFAWMLQFYVYKRFNFKKSWGCGLESKCKEKAALKVLOEYRKLGP 310
QY 309 MSCHIEIOMAIATAFAMIVLLITRKPFGVPGWSDLI-----NRKVGSGASGLSFIVLLIFALP 364
Db 311 LSFANINVLICFLVILWFSRDPGFMPCGLTVANWVEGETKYVSDATVAIFATLLFIVP 370
QY 365 TQYTFKYYCC-----GKGPTTAQIDAILSWEYVLRNIPMGLLFLGGGFALAVASRETGL 420
Db 371 SQKPKFNRSPQEEKSPVLI-APPPLDWKVTQEKVPMGIVLLGGGFALAKGSEASGL 429
QY 421 NIMISKAMQVILGNIVVQSTTFVLANFFSFAFNANVNVANIVILPILCEMSLALHLP 480
Db 430 SVMGMGQMEPLHAVPPAAITLISLLVAVFTECTSNVATTTFLPFIASMSRSIGLNFLY 489
QY 481 LITLPAICLGISMYVFLPVPSTPPNAIVTQYAHIKTKYFACCGIVPTTIIGISVALVNTWGL 540
Db 490 IMLPCTLSASFAMLPVATPPNAIVFTYGHKLVADMVKTGVIMNIIGVFCVFLAVNTWGR 549
QY 541 IIPFESKSPDDWA 553
Db 550 AIF-DLDHFPDWA 561

RESULT 7
US-10-173-519-2
; Sequence 2, Application US/10173519
; Publication No. US20020193582A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 69624, A Human Transporter Family Member
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: MP101-098P1RM
; CURRENT APPLICATION NUMBER: US/10/173,519
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/298,970
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-173-519-2
Query Match      30.2%; Score 875; DB 13; Length 568;
Best Local Similarity 35.7%; Pred. No. 1.1e-72;
Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;

QY 27 IIPLTILPILYGFQTDMAEF- KCLMLIVTMALLWITETLPIYVYVTFALPFLVFCPLGLVNA 85
Db 19 VTPLLLPLVI-----LMPAKFVRCAYVILMAIYVCTEVIPLAVTSLMPVLLPFLQILD 74
QY 86 ASIVCKQYFDTTIVFLGGLIVALGIEYSLHTRIALRVIRVGGSPRRFLVGLMSVTF 145
Db 75 SRQVCVQYKMDTNMLFLGGLIIVAVAVERNLHRIALRTLWVGAKPARMLMGFGVTAL 134
QY 146 MGLWISNSAGTAMMCPIVKALVNELDTN-----KIPPVYMTQEEEP-VVE 188
Db 135 LSMWISNTATTAMVPIVEAILQOMEATSAATEAGLELDVKGKAKELPGSQVIFEGPTLG 194
QY 189 GEPHPHSKITVAFYAGIAYASSIGGLTGLTGTLNLFVFRGIYTERPPTSTVETITANFM 248
Db 195 QEDQERKRLCKAMTLCICYAASIGGTATLTGTGNVVLGOMNELFPDSKOLVNEASWF 254
QY 249 FYSIPLMVINVTAVIAEL-ITHMGLFRPN-SKTGKIABANTNRKLMEDVLRQRIHID 307
Db 255 AFAPENMLV-----MLLFAWMLQFYVYKRFNFKKSWGCGLESKCKEKAALKVLOEYRKL 310
QY 308 PMSCHIEIOMAIATAFAMIVLLITRKPFGVPGWSDLI-----NRKVGSGASGLSFIVLLIFALP 363
Db 311 PLSFAINVLICFLVILWFSRDPGFMPCGLTVANWVEGETKYVSDATVAIFATLLFIV 370
QY 364 TQYTFKYYCC-----GKGPTTAQIDAILSWEYVLRNIPMGLLFLGGGFALAVASRET 418
Db 371 PSQKPKFNRSPQEEKTFFVP---PELLOWKVTQEKVPMGIVLLGGGFALAKGSEAS 427
QY 419 GLNIMISKAMQVILGNIVVQSTTFVLANFFSFAFNANVNVANIVILPILCEMSLALHLP 478
Db 428 GLSVMGMGQMEPLHAVPPAAITLISLLVAVFTECTSNVATTTFLPFIASMSRSIGLN 487
QY 479 LITLPAICLGISMYVFLPVPSTPPNAIVTQYAHIKTKYFACCGIVPTTIIGISVALVNTW 538
Db 488 LYIMLPCTLSASFAMLPVATPPNAIVFTYGHKLVADMVKTGVIMNIIGVFCVFLAVNTW 547
QY 539 GLIIPFESKSPDDWA 553
Db 548 GRAIF-DLDHFPDWA 561

RESULT 8
US-10-403-161-72
; Sequence 72, Application US/10403161
; Publication No. US20040043930A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-573C
; CURRENT APPLICATION NUMBER: US/10/403,161
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/370349
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384543
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/370969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/403748
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/372019
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/374379
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 05/779679
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/181045
; PRIOR FILING DATE: 2000-02-08
```

; PRIOR APPLICATION NUMBER: 10/055877  
; PRIOR FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: 60/262892  
; PRIOR FILING DATE: 2001-01-19  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 173  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 72  
; LENGTH: 568  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-403-161-72

Query Match 30.2%; Score 875; DB 15; Length 568;  
Best Local Similarity 35.7%; Pred. No. 1.1e-72;  
Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;  
QY 27 IIPILITPILYIGFQTDMAEF-KCLMLIVTMALLMITETPLPIYVTAFLPVLFCPLGLLVN 85  
DB 19 VTPLLLLPLVI-----LMPAKFVRCAYVILMAIYWCTEVIPLAVTSLMPVLLPFLQILD 74  
QY 86 ASIVCKQYFTDTTIVFVGLGILVALGIEYSNLHTRIALRVIRVGGSPRRFLVGLMSVSIF 145  
DB 75 SRQVCQYMKDNTMLFLGGLIIVAVAVERNLHRIALRTLLWVGAKEPARMLGFMGVTA 134  
QY 146 MGLWISNSAGTAMMCPIVKALVNELDN-----KIPPVYMTQEEP-VE 188  
DB 135 LSMWISNTATTAMVPIVEAILQOMEATSAATEAGLELVKGAKEPLGSGQVIFEGTIG 194  
QY 189 EGEPPHPSKIYAFYAGIYASSIGGLTIGTGNLVFRGIYTERPPTSTVITFANFM 248  
DB 195 QOEQERKRLCKAMTLCICVAASIGGTATLTGTGNVLLGQNNELFPDSKDLVNFASF 254  
QY 249 FYSIPLMVIVNVLIIAFILTHMGLFRPN-SKTGKIIAEANTNRKLMEDVLRORHIDLG 307  
DB 255 AFAPFNMLV-----MLLEFAMLVQVYMRNFKKSGCGLESKKNEAKALKVLOEYRKLG 310  
QY 308 PMSCHETOMIAFAFMIVLLITRKPGFVPGWSDLI-----NRKVGSASGLSFTVLILFAL 363  
DB 311 PLSPAEINVLICFLLVILWFSDPGFMGLTVAWVEGETKYVSDATVAIFVATLLFIV 370  
QY 364 PTQYTFKYKCC-----GKGPFTAQAIIDAILSWEYVLRNIPWGLLFLGGGFALAVASRET 418  
DB 371 PSQPKFNFRSQTEERKTPFP---PPLLDWKVTQEKVPWGIIVLLGGGFALAKGSEAS 427  
QY 419 GLNIMISKAMQVLIGLNPVQSIITFVLNFFSAFNANVVANIVLPILCEMSLALHLP 478  
DB 428 GLSVMMGQMEPLHVPVPAATLILSLVAVFTECTSNVATTLFLPFPASMSKSLGNP 487  
QY 479 LILTLPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTNTW 538  
DB 488 LYIMLPCTLSAFAMLPVATPPNAIVFTYGHKLVADMKVTGVMIMNIIGVFCVFLAVNTW 547  
QY 539 GLIIPPEKSPFDWA 553  
DB 548 GRAIF-DLDHFPDWA 561

RESULT 9  
US-10-490-080-1  
; Sequence 1, Application US/10490080  
; Publication No. US20040253597A1  
; GENERAL INFORMATION:  
; APPLICANT: Takeda Chemical Industries, Ltd.  
; TITLE OF INVENTION: Novel Protein and its DNA  
; FILE REFERENCE: P02-0109PCT  
; CURRENT APPLICATION NUMBER: US/10/490,080  
; CURRENT FILING DATE: 2004-03-17  
; PRIOR APPLICATION NUMBER: JP 2001-281992  
; PRIOR FILING DATE: 2001-09-17  
; PRIOR APPLICATION NUMBER: JP 2001-306873  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: JP 2002-113279

; PRIOR FILING DATE: 2002-04-16  
; NUMBER OF SEQ ID NOS: 42  
; SEQ ID NO 1  
; LENGTH: 568  
; TYPE: PRT  
; ORGANISM: Human  
US-10-490-080-1

Query Match 30.2%; Score 875; DB 16; Length 568;  
Best Local Similarity 35.7%; Pred. No. 1.1e-72;  
Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;  
QY 27 IIPILITPILYIGFQTDMAEF-KCLMLIVTMALLMITETPLPIYVTAFLPVLFCPLGLLVN 85  
DB 19 VTPLLLLPLVI-----LMPAKFVRCAYVILMAIYWCTEVIPLAVTSLMBVLLPFLQILD 74  
QY 86 ASIVCKQYFTDTTIVFVGLGILVALGIEYSNLHTRIALRVIRVGGSPRRFLVGLMSVSIF 145  
DB 75 SRQVCQYMKDNTMLFLGGLIIVAVAVERNLHRIALRTLLWVGAKEPARMLGFMGVTA 134  
QY 146 MGLWISNSAGTAMMCPIVKALVNELDN-----KIPPVYMTQEEP-VE 188  
DB 135 LSMWISNTATTAMVPIVEAILQOMEATSAATEAGLELVKGAKEPLGSGQVIFEGTIG 194  
QY 189 EGEPPHPSKIYAFYAGIYASSIGGLTIGTGNLVFRGIYTERPPTSTVITFANFM 248  
DB 195 QOEQERKRLCKAMTLCICVAASIGGTATLTGTGNVLLGQNNELFPDSKDLVNFASF 254  
QY 249 FYSIPLMVIVNVLIIAFILTHMGLFRPN-SKTGKIIAEANTNRKLMEDVLRORHIDLG 307  
DB 255 AFAPFNMLV-----MLLEFAMLVQVYMRNFKKSGCGLESKKNEAKALKVLOEYRKLG 310  
QY 308 PMSCHETOMIAFAFMIVLLITRKPGFVPGWSDLI-----NRKVGSASGLSFTVLILFAL 363  
DB 311 PLSPAEINVLICFLLVILWFSDPGFMGLTVAWVEGETKYVSDATVAIFVATLLFIV 370  
QY 364 PTQYTFKYKCC-----GKGPFTAQAIIDAILSWEYVLRNIPWGLLFLGGGFALAVASRET 418  
DB 371 PSQPKFNFRSQTEERKTPFP---PPLLDWKVTQEKVPWGIIVLLGGGFALAKGSEAS 427  
QY 419 GLNIMISKAMQVLIGLNPVQSIITFVLNFFSAFNANVVANIVLPILCEMSLALHLP 478  
DB 428 GLSVMMGQMEPLHVPVPAATLILSLVAVFTECTSNVATTLFLPFPASMSKSLGNP 487  
QY 479 LILTLPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTNTW 538  
DB 488 LYIMLPCTLSAFAMLPVATPPNAIVFTYGHKLVADMKVTGVMIMNIIGVFCVFLAVNTW 547  
QY 539 GLIIPPEKSPFDWA 553  
DB 548 GRAIF-DLDHFPDWA 561

RESULT 10  
US-10-718-359-6  
; Sequence 6, Application US/10718359  
; Publication No. US20050095240A1  
; GENERAL INFORMATION:  
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.  
; TITLE OF INVENTION: NaCT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION  
; FILE REFERENCE: 275.0008 0101  
; CURRENT APPLICATION NUMBER: US/10/718,359  
; CURRENT FILING DATE: 2003-11-20  
; PRIOR APPLICATION NUMBER: 60/428,469  
; PRIOR FILING DATE: 2002-11-22  
; PRIOR APPLICATION NUMBER: 60/459,441  
; PRIOR FILING DATE: 2003-04-01  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 6  
; LENGTH: 568  
; TYPE: PRT  
; ORGANISM: human NaCT



```

; ; CURRENT APPLICATION NUMBER: US-10/173,519
; ; CURRENT FILING DATE: 2002-06-17
; ; PRIOR APPLICATION NUMBER: 60/298,970
; ; PRIOR FILING DATE: 2001-06-18
; ; NUMBER OF SEQ ID NOS: 8
; ; SOFTWARE: FastSEQ for Windows Version 4.0
; ; SEQ ID NO 5
; ; LENGTH: 587
; ; TYPE: PRT
; ; ORGANISM: Rattus norvegicus
; ; US-10-173-519-5

```

Query Match	30.1%	Score 871	DB 13	Length 587
Best Local Similarity	37.4%	Pred. No. 2.8e-72		
Matches 213	Conservative 101	Mismatches 196	Indels 60	Gaps 17
Qy	30	LITLPIILYGFQTDMAEFKLLWLTVTWALLWITETLPIYVTALPPLVFCPLLLGLVNASIV	89	
Db	24	LLPLPLIV---QT--KEAYCAYSILMALLWCTEALPLAVTALFPVILFPLMGIMDASEV	78	
Qy	90	CKOYFTDTIVVFLGGLIIVALGIEYSNLHTRIALRIVIRVGGSPRRLFVGLMSYSTFMGLW	149	
Db	79	CIEYFKDNLIFVGGMLVAIAVEHNLHKRIALQVLLIIGVRPALLLGLFMLVTAFLSMW	138	
Qy	150	ISNSGAMTMCPIKALVNLDTNK-----IPVYMTQEEEPVEE-----CEP----	192	
Db	139	ISNTATTAMVPVIGHAVLELQSGSKDVEGNNNPTPELOEECPQEVTKLDNGQPVSAP	198	
Qy	193	--PHPSKITV--AFYAG----IAYASISGLGLTIGTGNLVFRGIVTERFPTSTVITTF	244	
Db	199	SEPTQKTQEHRSFQGLSLCICYSASIGGIATLTGTPNLVLQGVNSLFPQNGNVNF	258	
Qy	245	ANWFYSIPLMVINVTLVIAFL---ITHMGL-FRPNSTKGKIIAEANTNRKUME-DVL	299	
Db	259	ASWFGFAFPTMI-----LLAILAWLQVLFGVNRFRKNFGFG-----EGEERKQAAQVI	310	
Qy	300	RQRHIDLGPMSCHEIQMAIAFAFMIVLLITRKPGFVPGWSDLI--NRKVVGSASG---LS	354	
Db	311	KTORLLGPMFSFABKTVTLVFLVILVWLTREPGFFPGWGTVFANKEGQSMASDGTVAI	370	
Qy	355	FIVLLIFALPTQYTFYKCCG-----KGPFTAQDAIDAILSWEYVLRNIPWGLFLFLLGGF	409	
Db	371	FISLVMIIFPSKIP-----GLMQDPKKPGKLKAPPALLTKTVDNKKPMWNVILLGGGF	424	
Qy	410	ALAVASRETGLNIMISKAMQVLIGLPHNIVGSIITFVLANFPSAFNANVAVNIVLPILCE	469	
Db	425	ALAKGSQSGLSEWLGDKLTPLQHIHPPSATAVILCLLIAIETCSNVATTTLPILAS	484	
Qy	470	MSLALHLPLILTLTPACILGISMVFPLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIGIS	529	
Db	485	MAQAIHLHPLYVMLPCTLAASLAFMLPVATTPNAIVFSGGLKYSDMARAGFILNIIGVL	544	
Qy	530	VALVNTNTWGLIIFPSKSFDPDWAKEIKNQ	559	
Db	545	AITUSINSWSPIE-KLDTFPGSAHSNTSO	573	

```

RESULT 13
US-10-167-994-12
; Sequence 12, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIORITY APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12

```

```

; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-167-994-12

```

Query Match	30.1%	Score 871,	DB 14;	Length 587;
Best Local Similarity	37.4%;	Pred. No. 2.8e-72;		
Matches 213;	Conservative 101;	Mismatches 196;	Indels 60;	Gaps 17;
Qy	30	LITPLILYGFQDMAEFKLWLIVTMALIMWITETLPIYVTALFPLVFCFLLGLGNASIV	89	
Db	24	LLPLPLIV---QT--KEAYCAYGILMALLMCTEALPLAVTALFPVLFPIMGIMDASEV	78	
Qy	90	CKQYFTDTIVVFGGLIVAGIBYSNLHTRIAURIVRIVGGSRRRLRVGLMSTVSTFMGLW	149	
Db	79	CIEYFKDTNLLFVGGMLVAIAVEHNLHKRIALQVLLIIGVREPALLLGFMPLVTAFLSMW	138	
Qy	150	ISNSAGTAMCPVTKALVNELDNPK-----IPPVYMTQEEBPVE-----GEP----	192	
Db	139	ISNTATTAMVPVIGHAVLEQGSKKDVEGGNNPTFELQECPOKEVTKLDNGQPVSA	198	
Qy	193	--PHPSKITV--AFYAG-----IAYASSIGGLTGLGTGNLVRGITYTERPPTSTVEITF	244	
Db	199	SEPTQKTQEHHRFSQGLSLCICYSASIGGIATLTGTTPLVLQOGWSLFPQNGNVWF	258	
Qy	245	ANPMFYSIPLMWIWNVLVIIAEL-----ITHMGL--FRPNSTKGIIAEANTNRKLM-----DVL	299	
Db	259	ASWFGFAFPTMI-----LLLAWLWLQVFLGVNFRKNFGF-----EGEERKQAAPQVI	310	
Qy	300	RQRHIDLGPMSCHEIOWAIAFAFMVILLIIRKPGFVPGWSDLI--NRKVVSAGS-----LS	354	
Db	311	KTQRLGLGPNFSAEKTVTVLVLVLVWLNWTFREPGFPGWGDTPVANEKGQSMASDGTVAI	370	
Qy	355	FIYLLIFALPTQYTFKYCG-----KGPTAQAIADAILSWEXVLRNPWGLLFLILGGGF	409	
Db	371	FISLWFIIPSKIP-----GLMQDPKKPKKLKAPPAILLTKVTNDKQPNWIVILLGGGF	424	
Qy	410	ALAVASRETGLNIMI SKAMQVLLGLPNIVVQSITFVLNFFSAFNANVNVANIVLPILCE	469	
Db	425	ALAKGEQSGSLGMLGDKLTPLOHIPSATAVILCLLIAITFECTSNVATTTLLPLPILAS	484	
Qy	470	MSLALHLPLILTLPACLIGISMVYFIPVSTPPNAIVTQYAHIKTKYPACCGIVPTIIGIS	529	
Db	485	MAQAI CLHPYVWMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDARAGFLNIIIGVL	544	
Qy	530	VALVNTWNTGLIITFPESKSPFDMAKETNQ	559	
Db	545	AITLSINSWSIPIF-KLDTFFSVAHSNTSQ	573	

RESULT 14  
US-10-718-359-13  
; Sequence 13, Application US/10718359  
; Publication No. US20050095240A1

```

, GENERAL INFORMATION:
, APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.
, TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION
, FILE REFERENCE: 275.0008 0101
, CURRENT APPLICATION NUMBER: US/10/718,359
, CURRENT FILING DATE: 2003-11-20
, PRIOR APPLICATION NUMBER: 60/428,469
, PRIOR FILING DATE: 2002-11-22
, PRIOR APPLICATION NUMBER: 60/459,441
, PRIOR FILING DATE: 2003-04-01
, NUMBER OF SEQ ID NOS: 31
, SOFTWARE: PatentIn version 3.2
, SEQ ID NO 13
, LENGTH: 587
, TYPE: PRT
, ORGANISM: rat Nacd1

```

US-10-718-359-13

```

Query Match      30.1%; Score 871; DB 17; Length 587;
Best Local Similarity 37.4%; Pred. No. 2.8e-72;
Matches 213; Conservative 101; Mismatches 196; Indels 60; Gaps 17;

QY 30 LITLPIILYGFQTDMAEFKCLWLIVTWALLWITETLPIYVYATLPFPIFCPLLGLVNASIV 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 24 LLPLPLIV---QT--KEAYCAYSILMALLWCTEALPLAVTALFPVILFPLMGIMDASEV 78
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 90 CKQYFTDTIVVFLGGLVALGIEYSNLHTRIALVRIVGSPRRLFVGLMSVSTFMGLW 149
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 79 CIEYFKDTNLFVGLMVAIVEHNLKRIALQVLIIIVRPALLLLGLFMLVTAFTAFISMW 138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 150 INNSAGTAMMCPIKVALUNBELDTNKK-----IPVYVTOEEEPVE-----GEP 192
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 139 ISNTATTAMVPICHAVLELOQSGKQDVEGNNNPTFELQECPOKEVTKLDNGQPVSA 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 193 --PHPSKITV--AFYAG---IAYASSIGGLTGLGTNLVFRGIYTERPPTSTVEITF 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 199 SEPTQKTQEHHRPSQGLSLCICYASIGGIATLTGTPNLVQGVNSLFPQNGVNVF 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 245 ANWFYISIPLMVIVNVTLLIAFL---ITHMGL--FRNSKNTGKIIAEANTNRKLME--DVL 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 259 ASWFGFAFTMII-----LILLALWLVQLVFLGVNFRKNFGEG-----EGEERKQAAQVI 310
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 RORHIDLGPMSCHEIQMAIAFAMIVILLITRKGFVFGWSDLI--NRKVVGSAAG--LS 354
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 311 KTVRYLLGPMNSFAEKTIVLVLLVLMVLFREPFGFFGWDGTVPANEGKQSMASDGVIAI 370
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 355 FIVLLIFALPTQVTFKYCCG-----KGPTTAQIDAIDAILSWEYVLRNIPWGLLFLGGGF 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 371 FIVLNVFIIPSKIP-----GLMQDPKPKGLKAPPAILTWTKTVDNKWPNVILLGGGF 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 410 ALAVASRETGLNIMISKAMQVLGLPNIVQSGITFVLANFSAFNANVVANIVLPLICE 469
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 425 ALAKGSEQSGLSWIGDKUTPLOHPPSATAVILCLLIATFECTSNVATTTUFLPILAS 484
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 470 MSLALELHPILILTPACLGISWYFVLPVSTPPNNAIVTQYAHIKTKYPACCGIVPTIGIS 529
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 485 MAQAICLHPLYVMLPCTLAASLAFMLPVATPPNNAIVFSFGGLKVDNARAGFLNIGVL 544
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 530 VALVNTNTWGLIIPPSKSPFDWAKEIKNQ 559
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 545 AITLSINSWAPIF-KLDTPPSWAHSNTSQ 573
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-10-092-900A-46
; Sequence 46, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:29:45 ; Search time 16.7848 Seconds

(without alignments)  
3221.603 Million cell updates/sec

Title: US-10-017-479A-3

Perfect score: 2897

Sequence: 1 MAEPGEQRKVLGRCCIFHW.....PPESKSPFDNAKBIKNQTKI 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	877	30.3	568	JC7911	Na+-coupled citrat
2	830	28.7	593	I46528	sodium/dicarboxyla
3	747	25.8	599	T21613	hypothetical prote
4	725.5	25.0	545	T18694	hypothetical prote
5	687	23.7	595	A47714	Na+/sulfate cotran
6	672.5	23.2	526	S43561	YCR37C homolog K08
7	672.5	23.2	534	G88575	protein K08E5.2 li
8	642	22.2	539	A88546	protein R107.1 (im
9	617	21.3	520	C99880	hypothetical prote
10	610.5	21.1	539	S30871	hypothetical prote
11	545	18.8	552	F64546	sodium-dependent t
12	536	18.5	548	E71961	probable transport
13	442	15.3	446	F70302	transporter (pho87
14	436	15.1	461	I64080	probable membrane
15	433	14.9	462	B2510	transporter, Nadc
16	425	14.7	471	G81156	transporter, Nadc
17	425	14.7	471	A81948	probable transmemb
18	420.5	14.5	456	AF0093	Sodium, sulfate sym
19	396	13.7	487	C92215	probable transport
20	353	12.2	443	C69205	sodium/dicarboxyla
21	337.5	11.6	432	H64383	Na+ transporter -
22	330.5	11.4	478	A82336	probable transport
23	327.5	11.3	867	T40336	probable MSF trans
24	286	9.9	894	S45135	probable membrane
25	283.5	9.8	517	C98840	conserved hypothet
26	258.5	8.9	881	S46633	probable membrane
27	251	8.7	923	1 MMBY7C	hypothetical prote
28	239	8.2	513	2 I39534	probable membrane
29	223.5	7.7	450	2 A10909	probable membrane

30	214	7.4	589	2	H84072	sodium/sulfate sym
31	213	7.4	425	2	C75019	transport protein
32	208.5	7.2	610	2	A85870	probable transport
33	208.5	7.2	610	2	H91025	probable transport
34	208.5	7.2	610	2	B65001	probable transport
35	208	7.2	424	2	F71205	hypothetical prote
36	204.5	7.1	608	2	AC0798	probable sodium/su
37	191.5	6.6	428	2	G83803	hypothetical prote
38	189.5	6.5	612	2	S74936	sulfur deprivation
39	183	6.3	618	2	F84409	arsenite transport
40	182.5	6.3	429	2	G70528	probable arsa prot
41	174	6.0	157	2	S36784	mucin - rat (fragm
42	172.5	6.0	428	2	H70528	probable arsb prot
43	170	5.9	449	2	S76839	hypothetical prote
44	165.5	5.7	482	2	F71969	hypothetical prote
45	162.5	5.6	429	2	F87038	probable membrane
46	162	5.6	429	1	B41902	arsenical pump mem
47	161.5	5.6	425	2	F72315	conserved hypothet
48	161.5	5.6	431	2	A11599	arsenic efflux pun
49	161.5	5.6	487	2	B91122	hypothetical prote
50	161.5	5.6	487	2	A85967	hypothetical prote
51	160	5.5	430	2	G89962	hypothetical prote
52	159.5	5.5	610	2	AG0312	probable ion trans
53	158.5	5.5	487	1	Q8ECRS	Y9J4 protein - Esc
54	158	5.5	461	2	A82185	hypothetical prote
55	156.5	5.4	592	2	D95393	hypothetical prote
56	155	5.4	589	2	F75398	transporter, sodiu
57	154.5	5.3	431	2	AD1237	arsenic efflux pun
58	151.5	5.2	610	2	H83165	probable sodium/su
59	151	5.2	429	1	C41903	arsenical pump mem
60	150	5.2	487	1	B64795	ybds protein - Esc
61	150	5.2	487	2	G85560	probable membrane
62	150	5.2	487	2	C90710	probable membrane
63	149	5.1	436	2	D98100	conserved hypothet
64	148.5	5.1	436	2	G84024	arsenic efflux pun
65	145	5.0	421	2	G72300	conserved hypothet
66	145	5.0	436	2	C95236	membrane protein l
67	144.5	5.0	435	2	G69779	arsenical pump mem
68	142	4.9	368	2	C71131	hypothetical prote
69	141	4.9	445	2	AD2184	hypothetical prote
70	140.5	4.8	522	2	B64067	Na+/H+-exchanging
71	139.5	4.8	585	2	S69216	sulfur deprivation
72	137.5	4.7	496	2	C69903	proline permease h
73	137.5	4.7	501	2	AE0578	citrate carrier (i
74	137	4.7	512	2	B90050	hypothetical prote
75	135.5	4.7	469	2	T35526	probable integral
76	135	4.7	489	2	AF2970	conserved hypothet
77	135	4.7	489	2	D98312	hypothetical prote
78	134.5	4.6	477	2	G85579	probable membrane
79	134.5	4.6	477	2	F90728	probable dicarboxy
80	134.5	4.6	632	2	A71259	probable membrane
81	133.5	4.6	429	2	AE0406	arsenical pump mem
82	133	4.6	475	2	AI0010	probable membrane
83	132.5	4.6	477	1	B64813	ybhi protein - Esc
84	130.5	4.5	429	1	B25937	arsenical pump mem
85	130.5	4.5	590	2	AG3106	transporter Atu447
86	130.5	4.5	590	2	E98180	transporter, sodiu
87	130	4.5	801	2	A89862	Na+/H+ antiporter
88	128.5	4.4	522	2	A83804	involved in spore
89	126.5	4.4	684	2	B69308	conserved hypothet
90	126	4.3	471	2	F71543	probable dicarboxy
91	125	4.3	484	2	AE0887	probable membrane
92	124.5	4.3	417	2	AF2992	arsenical pump mem
93	124.5	4.3	417	2	D98291	probable arsenical
94	124.5	4.3	388	2	AC3587	sulfur deprivation
95	124	4.3	388	2	AF1541	antibiotic resista
96	123.5	4.3	440	2	G84069	hypothetical prote
97	123.5	4.3	473	2	B95376	probable Arc02 arg
98	123.5	4.3	700	2	H82428	iron(III) ABC tran
99	123	4.2	369	2	A75166	hypothetical prote
100	123	4.2	469	2	C90446	permease, multidru

## ALIGNMENTS

## RESULT 1

JC7911  
Na+-coupled citrate transporter NaCT - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Mar-2003 #sequence\_revision 31-Mar-2003 #text\_change 14-Jul-2003  
C;Accession: JC7911  
R;Inoue, K.; Zhuang, L.; Ganapathy, V.  
Biochem. Biophys. Res. Commun. 299, 465-471, 2002  
A;Title: Human Na+-coupled citrate transporter: Primary structure, genomic organization,  
A;Reference number: JC7911; MUID:22334959; PMID:12445824  
A;Accession: JC7911  
A;Molecule type: mRNA  
A;Residues: 1-568 <INO>  
C;Cross-references: GB:AY151833  
C;Comment: This transporter classified as a tricarboxylate transporter represents the first  
or the cellular entry of citrate by a process energized by the electrochemical Na+ gradient  
or the synthesis of fatty acid and chol 17p12-13  
C;Genetics:  
A;Gene: nact

Query Match	30.3%;	Score 877;	DB 2;	Length 568;
Best Local Similarity	35.7%;	Pred. No. 2.1e-59;		
Matches	198;	Conservative 99;	Mismatches 218;	Indels 40;
Gaps	10;			
Qy	27	IIPLITLPIIYGFOTDMAEF-KCLWLIVTMALLWITETLPIVVTALFPLVFCPLLGLVN	85	
Db	19	VTPLLLLPLVI-----LMPAKFVRCAVYIILMAIYWCTEVIPLEAVTSLMEVLLFPLFQILD	74	
Qy	86	ASIVCKQVPTDPIVVFGLGLVALGIEYSLNHLTRIALRIVIRIVGGSPRLFLVGLMSVSTF	145	
Db	75	SQVCVQYWKDTNMLFGLGLIVAVAVERNLHKEIALRTLWVGAKPARMLGFGCVTAL	134	
Qy	146	MGLWISNSAGTAMWCPIKVALVNLDTN-----KIPPYVMTQEEBPV-E	188	
Db	135	LSMWISNTATTAMVPIVEALIQOMEATSAATEAGLELVDKGRAKELPGSQVIFGPIGL	194	
Qy	189	EGEPPHPSKITVAFYAGIAYASSIGGLTGLTCTNLVFRGIVTERFPFTSTVEITPANFM	248	
Db	195	QEDQERKELCKAMTLCICYAASIGGTATLGTGPNVVLGQNNELFPDSDLVNFASWE	254	
Qy	249	FYSIPLMWIWNVLVIIAFLTHMGLFRPN-SKTKGIIIAEANTNRKLMEDVLRQIHIDIG	307	
Db	255	AFAPPNMLV-----MLLPFWMLQFYVFRNFPKSGGCGLSKKNEKAULKQVEERYKUG	310	
Qy	308	PMSCHIEQMAIAFAPMIVILLITRKEPGFVPGWSDLI-----NRKVVGASAGLSFTVLLIFAL	363	
Db	311	PLSFAEINVLCFELLVILWFSRDPGFMEGLTVAVWVEGETKYVSDATVAIVATLLFV	370	
Qy	364	PTQYTFKYCC-----GKGPTTAQDAIDAILSEWYVLRNIPKGLLFTLGGGFALAVASRET	418	
Db	371	PSQKPFNFRSQTEBERKTPFPV-----PPLLDWKVTQEKVPWGIVLLGGGFALUKGSEAS	427	
Qy	419	GLNIMISKAMQVLLIGLPIVTVQSIITFVLANFSAFNANVVVANIPLCILGEMSLALELHP	478	
Db	428	GLSVWVGKQMEPLHAVPPAAITLILSLVAVFTECTSNVATTTLFTLPIFASMSRSIGLNP	487	
Qy	479	LIILTLPACLGISNVYFLPSTPPNPAIVTQYAHIKTKYFACCGIVPTTIIGISVALVNTNW	538	
Db	488	LYTMLPCTLUSGAFAPMLPVATPPNAIVFTYGHKVDKVMKTVGMNIIGVFCVFLAVNTW	547	
Qy	539	GLIIFPEKSFDPWA	553	
Db	548	GRAIF-DLDHFPDWA	561	

RESULT 2  
146528  
sodium/dicarboxylate cotransporter - rabbit (fragment)  
C:Species: *Cryptotolagus cuniculus* (domestic rabbit)  
C:Date: 14-Feb-1997 #sequence revision 14-Feb-1997 #text change 09-Jul-2004

C:Accession: I46528

R;Pajor, A.M.  
J. Biol. Chem. 270, 5779-5785, 1995  
A;Title: Sequence and functional characterization of a renal sodium/dicarboxylate cotransporter  
A;Reference number: I46528; MUID:95197598; PMID:7890707  
A;Accession: I46528  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-593 <PAJ>  
A;Cross-references: UNIPROT:Q28615; EMBL:U12186; NID:g758383; PIDN:AAA99666.1; PID:g758383  
C;Superfamily: sodium/sulfate cotransporter

RESULT 3  
T21613  
hypothetical protein F31F6.6 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T21613

R.Percy, C.  
submitted to the EMBL Data Library, March 1996  
A:Reference number: Z19449  
A:Accession: T21613  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-599 <WIL>  
A:Cross-references: UNIPROT:Q93365; EMBL:Z69884; PIR:G11491  
A:Experimental source: clone F31F6  
C:Genetics:

A:Molecule type: DNA  
A:Residues: 1-599 <w1>  
A:Cross-references: UNIPROT:Q93655; EMBL:Z59884; PIDN:CAA93752.1; GSPDB:GN00028; CESP:F316  
A:Experimental source: clone F31P6  
C:Genetics:

A:Gene: CESP.F31F6.6  
A:Map position: X  
A:Introns: 39/3; 142/2; 207/2; 255/2; 285/2; 397/1; 437/3; 493/2; 537/3  
C:Superfamily: sodium/sulfate cotransporter

Query Match 25.8%; Score 747; DB 2; Length 599;  
Best Local Similarity 30.1%; Pred. No. 2e-49;  
Matches 174; Conservative 114; Mismatches 223; Indels 68; Gaps 10;

Qy 23 KASIII---PLITLPIYGFQDMAEFKCLWLIIVTMALLMITETLPIYVTFALFVFCP 79  
Db 15 KQSFVIMGALLIFSPLLMFVGDHGLQAKCLYCVAVMGVYVFEALPLAIFAPIMILFP 74

Qy 80 LGLVNASIVCKQYPTDITVIFLGLGLIIVGALGIEYNLHTRIALRVIRVGGSPRLFVGL 139  
Db 75 LFGIMRSEVAYALPDCTFLFMGLMVALAVEKCELHARVALFVLKTVGSEPARVMAGF 134

Qy 140 MSVSTFMGLWISNSAGTAMWCPIVKALVNEIDTN-----KIPFVYMTQ 182  
Db 135 MGVTFGLSMISNTATTALMPILQSVITELVSNHRMEDVALCEAHNSRKHSVGMRR 194

Qy 183 EEPVEEGE-----PHPSKITVAFVAGIAYASSIGLGLTIGTNLVRGIYT 232  
Db 195 LSLPNNENIKREENDTAMSPREQMAKGLMLSVCFSANIGGAATITGASNLVLGQLN 254

Qy 233 ERFPSTVEITFANPMFYSIPLMIVNVTVLIIAFLITHMGLFRPNSTKGIIEANTR 292  
Db 255 ELFPGADTGVNLFSLWIFAFPMVFCCLIVCWVLYL-----LYLRDAPKGSII---VTR 305

Qy 293 KLMDVLQRHIDLGPMSCHEIQMAIAFAMVILLITRKPGFVPCWSDL----- 341  
Db 306 K-----LQKNELHAFSAEMAVIFCFALLILWILREPQVPGWGMFKDELVFKSLT 360

Qy 342 -----INRKVGSASGLSFVLLFALPTQYTFPKYCGKGP-ATAIDAISWEVYL 394  
Db 361 EKONTHLTFRVSDATSAFVILFLTPE-----KLPSSRGSSBQRKASGLLDWATQ 415

Qy 395 RNIPMGLLFLGGGFALAVASRETGLNIMISKAMQVILGLPNVQSTFVLANFFSAFN 454  
Db 416 DRFPWSVFLGGGFALAAVKGESGLSHDIGAIMRYLDVFNHNIIMLCIIISVLTNVC 475

Qy 455 ANVVANIVLPILCMSIALELHPILITPLACLGISMVYFIPVSTPPNAIVTQYAHIKTK 514  
Db 476 SNTVIASIFIPVABELARSLEIDPLNFMPLVTISASPAFLPLPVPAPNAIVFSSGYLKV 535

Qy 515 YFACCGIVPTIIGISVALVNTWGLIIFPSSKSPDWA 553  
Db 536 DMFVSGLCVTILGCVVLSMLNMLLWAGFVF-NLHLFPQWA 573

RESULT 4  
Tl8694  
hypochemical protein B0285.6 - Caenorhabditis elegans  
A:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
A:Accession: Tl8694  
R:Submitted to the EMBL Data Library, June 1994  
A:Reference number: Z19007  
A:Status: preliminary  
A:Accession: Tl8694  
A:Molecule type: DNA  
A:Residues: 1-545 <WIL>  
A:Cross-references: EMBL:Z34533; PIDN:CAA84299.1; GSPDB:GN000021; CESP:B0285.6  
A:Experimental source: clone B0285  
C:Genetics:  
A:Gene: CESP:B0285.6  
A:Map position: 3  
A:Introns: 6/3; 56/3; 95/2; 107/2; 160/2; 175/1; 211/2; 287/2; 353/1; 393/3; 494/3  
C:Superfamily: sodium/sulfate cotransporter

Query Match 25.0%; Score 725.5; DB 2; Length 545;  
Best Local Similarity 31.5%; Pred. No. 7.9e-48;

Matches 174; Conservative 114; Mismatches 212; Indels 53; Gaps 13;  
Qy 43 DMAEFKCLWLIIVTMALLMITETLPIYVTFALFVFCPCLGLVNASIVCKQYFTDITVIFL 102  
Db 4 ESTEGKCAFVILITWCSYVVAEWPFLAVTSFIPMIALPFLGIVSIVKAVPKYFADTNIVFF 63

Qy 103 GGLVAGIEYNLHTRIALRVIRVGGSPRLFVGLMSVSTFMGLWISNSAGTAMWCPI 162  
Db 64 NSLMSLAVBECQLHRIALQMLTYVGTTRPHLMAGFMIIITPSISLWISDACCALMAPI 123

Qy 163 VKALV-----NELDTNKIFPVYMTQEEBEVE-EGBPHPHSKITV----- 200  
Db 124 AYALLEIMIPKMPKEKENEIEVKIF-----DKEDPEEKEKKLDTLSRVSRRDRGICK 178

Qy 201 AFYAGIAYASSIGLGLTIGTNLVRGIYTERFPTSTVITFANPMFYSIPLMIVNV 260  
Db 179 CMLLVAHASLIGGTGTINSTGNLIIFRDNIETKPNEDHGI SYLSWMAFAPMIFMYF 238

Qy 261 T---LVIIAFLITH--MGLFRPNSTKGIIEANTRKLMEDVLQRHI-----DLGPM 310  
Db 239 SSWFIVQLQFLGPRHLMGMPREPTETEK-----QEEVAKRAVWKSVDQLGPM 287

Qy 311 CHEIQMAIAFAFMVILLITRKPGFVPCWSDLINR-KVVGSAAGLSFVILLIPALPTQYTF 369  
Db 288 WAEXSTLIVFVLVLSWSSDPKVPICWSDLFRGYVTDSCGL-VAVLFLFIPWKKPD 346

Qy 370 FKYCGKGPFTAQDAIDAILSWEYVLRNIPWGLLGLGGFALAVASRETGLNIMISKAM 429  
Db 347 PRIFRKDSRPSVQEPDLWDVCVRRFPWSIILLGLGAFISDAVRVSGLSLIACSLN 406

Qy 430 VLIG-LPNIIVQSTFVLANFFSAFNANVVANIVLPILCMSIALELHPILITPLACIG 488  
Db 407 STISKMPFFVQIILSIWVVMTEFSTNSATASIFIPISFKMAEAVGAHPLFYSIPTAIG 466

Qy 489 ISMVYFIPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTWGLIIFPESKS 548  
Db 467 PSFSLPMPATPANAIYETKTRIMDMVSGVFLNFCIAITAINMTWAFWLF-NMGT 525

Qy 549 FPDWA-KEIKNOT 560  
Db 526 YPDVALRHATNMT 538

RESULT 5  
A47714  
Na+/sulfate cotransporter, renal - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Jun-1994 #sequence\_revision 02-Jun-1994 #text\_change 09-Jul-2004  
C:Accession: A47714  
R:Markovich, D.; Forgo, J.; Stange, G.; Biber, J.; Murer, H.  
Proc. Natl. Acad. Sci. U.S.A. 90, 8073-8077, 1993  
A:Title: Expression cloning of rat renal Na+/SO4(2-) cotransport.  
A:Reference number: A47714; MUID:93376745; PMID:7690140  
A:Accession: A47714  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-595 <VAR>  
A:Cross-references: UNIPROT:Q07782; GB:L19102; NID:g310182; PIDN:AAA41677.1; PID:g310183  
C:Superfamily: sodium/sulfate cotransporter

Query Match 23.7%; Score 687; DB 2; Length 595;  
Best Local Similarity 28.9%; Pred. No. 7.7e-45;  
Matches 173; Conservative 116; Mismatches 220; Indels 90; Gaps 14;

Qy 27 IIPILITPLIYGFQDMAEFKCLWLIIVTMALLMITETLPIYVTFALFVFCPLGLVNA 86  
Db 20 VLVLPLPLIIRS-----KEACAYILFVIAETWTEALPLSITALLPGLMFPFMGIMSS 74

Qy 87 SIYCKQYFTDITVIFLGLGLIIVGALGIEYNLHTRIALRVIRVGGSPRLFVGLMSVST 146  
Db 75 THVASAYFKDFHLLILGIVCLATSIIEKNLHRIALRMVMMVGNVPAWLTIGFMSSTAF 134

Qy 147 GLWISNSAGTAMWCPIVKALVNEIDTN-----KIPFVYMTQEEEP 186



```

Db      425 NICSNTVTASIEVPIVATLQAGHHPFTMLPTTLASSFAFIFVPGTPPNAIVFGSGMV 484
QY      512 KTKYFACCGIVPTIIGISVALVNTWGLIIPPEKSPDPDAKEIKNOT 560
Db      485 KVSDFAFVGGIISLELLVLTVMYSIAVLTLP-LLEPFTWA-IIANST 531

RESULT 8
A88546
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
R:Accession: A88546
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; PMID:9906913; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: A88546
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-539 <STO>
A:Cross-references: UNIPROT:P32739; GB:chr_III; PIDN:CAA78468.1; PID:g3879033; GSPDB:GNC
A:Note: predicted using Genefinder
C:Genetics:
A:Gene: R107.1
A:Map position: 3
C:Superfamily: sodium/sulfate cotransporter

Query Match      22.2%; Score 642; DB 2; Length 539;
Best Local Similarity 29.6%; Pred. No. 1.9e-41;
Matches 175; Conservative 109; Mismatches 209; Indels 98; Gaps 20;

QY      3  EPGEORKEVLCRCIFHWGKASIII--PLITPLIYVGFQTDMAEFKCLMLIVTMALLW 60
Db      2  KPSPQRTLI-----KCLLVLLGLVAVPLLFEG-----PEYRCLFSIIPLSTYW 45
QY      61  ITETPLIVVTALFPLVFCPLGLVNASIVCKQYFTDTIIVFLGLIIVAGLIEYNSLHTRI 120
Db      46  IGEAFPIGVTSFLPLALYPIQIVPSKQISPVYFQDSIVLFWCTLIIMAWAEATGLHRI 105
QY      121  ALRVIRIVGSGPRRLFLVGLMSVSTFMGLWISNSAGTAMCPIVKALVNLDELTKNIFPVYM 180
Db      106  ALKLLTKVGAK-----QPVSFF-----VSDTACTALMCTPAVALLMSMSD-----AV 147
QY      181  TQEEPEVEEGEPHPSKITV-----AFVAGIAYASSIGGLGLTIGTG 222
Db      148  OHLKEDHRKPKPP-PDQATVAEKLRIIDMTPODAGFCFCKALILACAHASLIGGTAIITSG 206
QY      223  TNLVFRGIYTEREPTSTVEITFANFMFYSIPLMIVNVT--LVIAFLI---THMGLFRP 277
Db      207  PNLVFRENIHKRYEGQVTMYLQMWFAIDPMFVYLLASVIIIVCYPMGSGTFAIRFER 266
QY      278  NSKTGKIIAEANTNRKLMEDVLRQRIIDLGPMSCHIEIQMAIAFAFMIVLLITRKPGFVPG 337
Db      267  PSK-----EEAHL-KKLIENKIQTMYEDLDGVSNGEKSVFVFILLIGSWISRPDGTTPG 320
QY      338  WSDLINRK--VVGASGLSFVLLIPALPTQYTFPKYCCGKGPPTAQAIID---AIIISWEY 392
Db      321  WGDLLPHRNFTSDSVSGV-LISCILFVMP-----KDPE--DPIDPMADILKWTGD 366
QY      393  VLNRIPWGLLELLGGFPALAVASRETGLNTIMISKAMQ-VLIGLNVVQSIITFLVANFFS 451
Db      367  MKSKFWSCTLLIGAGVAISGVDKSLRSLICGMKNIFVGMSSLPQLFTVTTIIVIMT 426
QY      452  AFANANVVANIVLPILCMSLALHLLPLITLPACLGISMVYFLPVSTPPNAIVTQVAHI 511
Db      427  EFASNVSTGSIFFIISLGVAESMGVHPLYALPTTVACSFAPMLPISTPPNAVVDYTKVI 486
QY      512  KTKYFACCGIVPTIIGISVALVNTW-----GLIIFP-----ESKSPF 550

```

```

Db      487  SMVEMIVCGFLLNIACILITSLANNTWTYFIFSLNIPFENIVIVISENSYSP 537

RESULT 9
C89980
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
R:Accession: C89980
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: C89980
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-520 <KUR>
A:Cross-references: UNIPROT:Q99SX1; GB:BA000018; PID:g13701709; PIDN:BAB43002.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA1732
C:Superfamily: sodium/sulfate cotransporter

Query Match      21.3%; Score 617; DB 2; Length 520;
Best Local Similarity 29.4%; Pred. No. 1.5e-39;
Matches 155; Conservative 102; Mismatches 197; Indels 74; Gaps 13;

QY      26  IIPILITPLIYVGFQTDMAEFKCLMLIVTM--ALLMITETPLIYVVTALFPLVFCPLLGL 83
Db      37  ILGPLLELLTLLFFHPQDLPWKGVYVLAITLWIAITWITEAIPAAATSLPLVLLPLGHI 96
QY      84  VNASIVCKQYFTDTIIVFLGLIIVAGLIEYNSLHTRIALRVIRIVGSGPRRLFLVGLMSVS 143
Db      97  LTPEQVSEYGNDIIFLGLGFILAIAMRNWNLHTRVALTIINIGASTSKILLGFMVAT 156
QY      144  TFMGLWISNSAGTAMCPIVKALVNLDELTKNIFPVYMTQEEPEVEEGEPHPS--KITVA 201
Db      157  GFLSMFVSNAAVMIMIPIGLAI-----KEAHDQEAANTNQTISOKEFS 202
QY      202  FYAGIAYASSIGGLTIGTGNLVFRGIYTERPPTSTVEITFANFMFYSIP-LMIVNV 260
Db      203  LVLAIGVAGTIGGLTIGTTPPLIILKGQYMQHFGH---EISFAKMMIVGIPTVIVLLGI 259
QY      261  TLVIAFL-ITHMGLFRPNSKTGKIIAEANTNRKLMEDVLRQRIIDLGPMSCHIEIQMAIA 319
Db      260  TWLYLRVAVFRHDLUKYLPGGQT-----LIQKLDLGLKMKYEEKVVOQT 303
QY      320  FAFMIVLLITRKPGFVPGNSDLINRKVVGSGLSFTVLITFALPTQVTFPKYCCGKGP 379
Db      304  FVLASLLWITRE-PLKKWE--VTSSVADGTIAI-FISILLFIIPAKNT----- 348
QY      380  TQAIDAILSWEVYLRNIPWGLLFLGGFPALAVASRETGLNTIMISKAMQVILGLPNIW 439
Db      349  --EKHRIIDWE-VAKELPWGLVILFGGLALAKGISEGLAKWLGQLKSLNGVSPILI 405
QY      440  QSITFVLNFFSAFNANVVANIVLPILCMSLALHLLPLITLPACLGISMVYFLPVST 499
Db      406  VIVITIFVLTEVTSNTATATMILFILATLSVAVGVHPLLMLLMAAANCAVMPLPVGT 465
QY      500  PPNAIVTQVAHIKTKYFACCG-----IVPTIIGISV 530
Db      466  PPNAIFGSGKISKQWASGVFWFNLSAIIIIIVVYVMPVIVLGIDI 513

RESULT 10
S30871
C:Species: Caenorhabditis elegans
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
R:Accession: S30871
R:Thomas, K.

```

submitted to the EMBL Data Library, July 1992

A;Reference number: S30871  
A;Accession: S30871  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-539 <THO>  
A;Cross-references: EMBL:Z14092  
C;Genetics:  
A;Genetic code: SGC4  
A;Introns: 30/3; 80/3; 119/2; 217/2; 251/3; 293/2; 390/3; 416/1; 445/3; 491/3  
C;Superfamily: sodium/sulfate cotransporter  
C;Keywords: transmembrane protein

Query Match 21.1%; Score 610.5; DB 2; Length 539;  
Best Local Similarity 29.3%; Pred. No. 4.9e-396;  
Matches 165; Conservative 107; Mismatches 206; Indels 85; Gaps 18;

QY 29 PLITLPIIYGFQDMAEFKCLMIVTMALLWITETLPIYVITAFPLVFCPLGLVNASI 88  
DB 19 PLVAVPLLPFG-----PEYCLFSMFLSTYIWEAFPIGVTSFLPALYPIILQVPSKQ 73  
QY 89 VKQYFTDTIVFVGLGIVAGLEYNLHTRIALRVIRVGGSPRLFFVGLMSVSTFMGL 148  
DB 74 ISPVYFKDSVLFMCTLINAMAVEATGLHRSIALKLLTKVGAK-----QPVSPF--- 122

QY 149 WINSAGTAMKCIPIKALVNELDTNKIFPVYMTQEEPEVEGEPHPSPKITT--- 200  
DB 123 -VSDTACTALMCTPAVALMSMD-----AVQHLKEDHSKPKPP-PDDATVAEKLMSMD 174

QY 201 -----AFVAGIAYASSIGLGLTIGTGNLVFRGIYTERPPTSTVEITFANMFY 250  
DB 175 MTPQDAGFCFALILACAHASLIGTALITSTGNLVFRNIHKSYPEGQVTMYLQMWVF 234

QY 251 SIPLMIVNVT--LVIIAFIL-----THMGLPRPSKTKGKIIAEANTWRKLMEDVLRORHID 305  
DB 235 AMPMPFVYLLASIIILVCYFMGSPSTFARWFESPCK-----EEAHL-KKLEKNIQTMED 288

QY 306 LGPMSCHEIOMIAFAFMIVLITRKPGFVPGVDSGLI--NRKVGSASGLSFIVLLIFAL 363  
DB 289 LGDVSGEKSVFVFILLIGSWISRDGFTPGWGDLLPHSNFMSDSVSGV-LISCLIFVW 347

QY 364 PTQYTFPKYCGKGPPTAQAIID---AIIWSWEYVLRNIPWGLLFLGGSFALAVASRETGL 420  
DB 348 P-----KDPF--DPIDPMAPILKWTDMKSKFSWSTLLIGAGYAIISGVGDKSGL 394

QY 421 NIMISKAMQ-VLIGLPIVIVQSTIFVLANFFSAFNANVVANIVLIPILCEMSLALHLPL 479  
DB 395 SSLISCGMKNI FVGMSSLPQLQTLVTTIIVVMTEFASNVSTGSIPIPIISLGVASSMGVHPL 454

QY 480 ILTLPACGLISWYVFLPVSTPPNAIVTQYAHIKTKYPACCGIVPTIIGISVALVNTNTW- 538  
DB 455 YLALPTTVACSFAPMLPISTPPNNAVVDTKVISVEMVMVGFLNLNACMLTSLNMTWT 514

QY 539 ----GLIIFP-----EKSFPF 550  
DB 515 YFIFSLNIFPENIVMSSENSYP 537

RESULT 11  
F64546  
sodium-dependent transporter - Helicobacter pylori (strain 26695)  
C;Species: Helicobacter pylori  
C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C;Accession: F64546

R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
Nature 388, 539-547, 1997  
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A;Reference number: A64520; MUID:97394467; PMID:9252185  
A;Accession: F64546  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-552 <TOM>  
A;Cross-references: UNIPROT:O25003; GB:AE000541; GB:AE000511; NID:g2313299; PIDN:AA00728;  
C;Superfamily: sodium/sulfate cotransporter

Query Match 18.8%; Score 545; DB 2; Length 552;  
Best Local Similarity 29.5%; Pred. No. 5.2e-34;  
Matches 156; Conservative 89; Mismatches 190; Indels 94; Gaps 16;

QY 54 VTMALLWITETLPIYVITAFPLVFCPLGLVNASIVCKQYFTDTIVFVGLGIVAGLEY 113  
DB 72 VLMGIWMTTEADLPATALLPLVFSVSDQFASVSSSYASPIIFLPMGGFILLALSMOK 131

QY 114 SNLHTRIALRVIRIVGGSPRLFFVGLMSVSTFMGLWISNAGTAMMCP1-----VKAL 166  
DB 132 WNLHTRIALSIILLVGTSPRLLIGFMATGFLSMWNTATFVVMPLVGMVSLQVAKL 191

QY 167 VNELDTNKIFPVYMTQEEE-----PVEGEPHPSPKITTVAFYAG 205  
DB 192 VGKEDASNSW-----HQKEBITKAHGGIMSNIVHKGKIDITQVIOEKTITTYRTNFSICLMIG 247

QY 206 IAYASSIGLGLTIGTGNLVFRGIYTERPPTSTVEITFANMFYSIPLMWIVNVTVII 265  
DB 248 IAYASSIGLGLTIGTGNLVFRGIYTERPPTSTVEITFANMFYSIPLMWIVNVTVII 300

QY 266 AFLITHMGLFRPNSKT---GKIIAEANTNRKLMEDVLRQRHIDLGPMSCHEIOMIAF-- 320  
DB 301 AWLLTYVIFPLKIKEIPGK-----EVIRVELKKLGRLSOAEISVGIIFIL 347

QY 321 -----AFMVLITRKPGFVPCWSDLNKRKVVGSAGLSFIVLLIFALPTQYTFPKYCGG 375  
DB 348 ASLGMIFLGMVL-----KSGVKIDKIDSIVAMGVS---ALLFILPANH----- 388

QY 376 KGPFTAQAIIDAILSWEYVLRNIPWGLLFLGGSFALAVASRETGLNIMISKAMQVLIGLP 435  
DB 389 QG-----DRLLDNG-VAKLPWDVLLFLGGLLALSQFSKTGLSLWHLVSGSFHLP 440

QY 436 NIWQSTIFVLZANFFSAFNANVVANIVLPIIL--CEMSLALSLH-PLIITLPACLGISMV 492  
DB 441 ILPIIWMVLMVIFLITETITNTATAAFLPVICGVAMGMYENHOSLLTIPVALSATCA 500

QY 493 YELPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTNTWGLI 541  
DB 501 FMLPVVTPPNAIAYGGYVKITDMIKAGLWMLNVG--VVLITFSYFLV 547

RESULT 12  
E71961  
probable transporter - Helicobacter pylori (strain J99)  
C;Species: Helicobacter pylori  
A;Variety: Strain J99  
C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C;Accession: E71961

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
Ives, C.; Gibson, R.; Merzberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;  
Nature 397, 176-180, 1999  
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A;Reference number: A71800; MUID:99120557; PMID:9923682  
A;Accession: E71961  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-548 <ARN>  
A;Cross-references: UNIPROT:Q9ZML8; GB:AE001458; GB:AE001439; NID:g4154713; PIDN:AA005784;  
A;Experimental source: strain J99  
C;Genetics:  
A;Gene: jhp0200  
C;Superfamily: sodium/sulfate cotransporter

Query Match 18.5%; Score 536; DB 2; Length 548;  
Best Local Similarity 29.0%; Pred. No. 2.5e-33;  
Matches 151; Conservative 96; Mismatches 197; Indels 76; Gaps 17;

QY 54 VTMALLWITETLPIYVITAFPLVFCPLGLVNASIVCKQYFTDTIVFVGLGIVAGLEY 113



Qy	397	IPNGLIFLGGGFALAVASRETGINIMISKAMQVLI-GLPNIVVQSIITFVLNFFSAFNA	455	
Db	316	TDWGVLMFGGGJTLTSAVLKDSGASKILADSIYFMDGGHFYLIIGLVLAAFIIFLTFEFTS	375	
Qy	456	NVVVANIVLPILCEMSLALHPLILITLPACLGISVMVFLPVPSTPPNAIVTQVAHIKTKY	515	
Db	376	NTASALLVIFITISIAQSLGMPGILGALLIGIGASCAPMLPVATPPNAIVFGSGQVKQSE	435	
Qy	516	FACCGIVPTTIIGISV	530	
Db	436	MVKVGFLLNLVCVVV	450	
RESULT 15				
B82510				
transporter, NadC family VCA0025 [imported] - Vibrio cholerae (strain N16961 serogroup O				
C:Species: Vibrio cholerae				
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004				
C:Accession: B82510				
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;				
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.				
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.				
Nature 406, 477-483, 2000				
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.				
A:Reference number: A82035; MUID:20406833; PMID:10952301				
A:Accession: B82510				
A>Status: Preliminary				
A:Molecule type: DNA				
A:Residues: 1-462 <HEI>				
A:Cross-references: UNIPROT:Q9KNE0; GB:AB004346; GB:AB003853; NID:9657401; PIDN:AAF9593				
A:Experimental source: serogroup O1; strain N16961; biotype El Tor				
C:Genetics:				
A:Gene: VCA0025				
A:Map position: 2				
C;Superfamily: probable transporter MJ0672				
Query Match 14.9%; Score 433; DB 2; Length 462;				
Best Local Similarity 23.8%; Pred. No. 1.6e-25;				
Matches 124; Conservative 96; Mismatches 170; Indels 132; Gaps 14;				
Qy	43	DMAEFKCLW-----LIVTWALLWITETLPPIVYTALPPLVFCPLLGLVNASI	88	
Db	28	DVALFLALYHLPPEHNVLGISMLAFIAVLMLTEALHVTITAILVPVMAVFFGIFETQA	87	
Qy	89	VCKQYFTDTIVVFLGGLIVALGIYSNLHTRIALRIVIGSSPRRLFVGLMSVSTFMGL	148	
Db	88	ALNNFANSIIIFELGGFALAAHHQGLDKVIADKVLAMAGCKMSVAVFMLFGVTALLSM	147	
Qy	149	WISNSAGTAMCPIVKALVNELDNTNKIPPVYMTQEBEPVEEGEPHPHPSKITVAFYAGIAY	208	
Db	148	WISNTATAAMMLPLVLGVLSKVDADKQRTY-----VFVLLGVAY	187	
Qy	209	ASSIGGLGTILGTGNLVFRGIYTERPFTSTVEITFANFMFYSIP-LMVIIVNVTLVIIAF	267	
Db	188	SASIGGIATLVGSPENAI-----AAAEVGLSFTDMKFKGLPTAMMMLPMAIAILYF	238	
Qy	268	LI--THMGLER-----PNSKTGKIABEANTNRKLMEDVLRQRHIDLGPMSCHETQMAIAF	320	
Db	239	LKPTLNGMFELDRAPVNDKGVVT-----LG-----IF	268	
Qy	321	AFMIVLLITRKFGFVPGNSDLNRKVGSA-----GLSFVILLIFALPTQYTFKYCCG	375	
Db	269	GLTVFLWIFSSP-----INAAFGFKSFDTLVALGAILMLSA-----	306	
Qy	376	KGPFTAQIDAILLSWEYVLRNIPWGLLFLGGGFALAVASRETGINIMISKAMQVLI---	432	
Db	307	-----RVVHWKEIQKTADWGVLLLFGGGLCLSNVLKQTGTSVFLANALSDMVSHM	356	
Qy	433	GLPNIVVQSIITFVLNFFSAFNAVNVANIVLPILCEMSLALHPLILITLPACLGISMV	492	
Db	357	GIFVILVVAITVV--FLTEFASNTASALLIPVATVAEAFGMSPVLLSVLIAVAASCA	414	
Qy	493	YFLPVPSTPPNAIVTQVAHIKTK-----YFACCGIVPTI	525	

Db

415 FMLPVATPPNAIVPASGHIKQSEMMRVGLYLNIAICIGLLTAI

456

Search completed: June 30, 2005, 09:00:15

Job time : 18.7848 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:39 ; Search time 75.6279 Seconds  
(without alignments)  
3805.323 Million cell updates/sec

Title: US-10-017-479A-3

Perfect score: 2897  
Sequence: 1 MAEPGEQRKFLVGRCCIFHW.....FPESKSFDPWAKEIKNQTKI 562

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Listing first 100 summaries

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt\_03.\*

1: uniprot\_eprot.\*  
2: uniprot\_trembl.\*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2897	100.0	562	1	IND2_DROME
2	1543	53.3	505	2	Q811B6
3	1494.5	51.6	504	2	Q86B89
4	1382.5	47.7	572	1	IND1_DROME
5	1382.5	47.7	590	2	Q7KUS6
6	1040	35.9	896	2	Q7QIT2
7	890	30.7	605	2	Q666R0
8	880.5	30.4	586	2	Q6AZR9
9	875	30.2	568	2	Q86YT5
10	871	30.1	568	2	Q6ZMG1
11	871	30.1	587	2	Q35055
12	868.5	30.0	586	1	S132_MOUSE
13	855.5	29.5	572	2	Q8CJ44
14	852.5	29.4	622	2	Q57661
15	851.5	29.4	624	2	Q642N5
16	846	29.2	613	2	Q803K7
17	844	29.1	592	1	S132_HUMAN
18	843	29.1	602	2	Q6DUN7
19	842	29.1	613	2	Q6P3L2
20	840	29.0	600	1	S133_MOUSE
21	839	29.0	602	2	Q6E7G8
22	833	28.8	600	1	S133_RAT
23	832.5	28.7	602	2	Q81VB1
24	831.5	28.7	602	1	S133_HUMAN
25	830	28.7	593	1	S132_RABIT
26	818	28.2	414	2	Q7PF67
27	814.5	28.1	572	2	Q67BT3
28	808	27.9	587	1	S132_RAT
29	789	27.2	601	2	Q9W7I2
30	781	27.0	596	2	Q6NMY4
31	764	26.4	625	2	Q8BZ82

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

32	763.5	26.4	582	1	NAD1_CAEEL
33	756	26.1	583	2	Q6NY28
34	754	26.0	583	2	Q6PE27
35	741	25.6	626	2	Q8N631
36	738.5	25.5	627	1	S134_HUMAN
37	724.5	25.0	577	1	YKG6_CAEEL
38	698.5	24.1	601	2	Q7ZWL2
39	698	24.1	520	2	Q8N8K4
40	696	24.0	595	1	S131_HUMAN
41	687	23.7	595	1	S131_RAT
42	680.5	23.5	594	1	S131_MOUSE
43	680.5	23.5	594	2	Q80YB5
44	677	23.4	551	1	NAD2_CAEEL
45	676	23.3	619	2	Q6L970
46	673.5	23.2	566	1	NAD3_CAEEL
47	639	22.1	546	2	Q65NCO
48	618	21.3	520	2	Q8NVSS
49	618	21.3	520	2	Q6G816
50	618	21.3	520	2	Q6GFE0
51	617	21.3	520	2	Q99SX1
52	617	21.3	520	2	Q7A4P8
53	580	20.0	495	2	Q67T81
54	572	19.7	548	2	Q6ZFH7
55	571	19.7	552	2	Q8NE33
56	551	19.0	510	2	Q8NTS7
57	551	19.0	527	2	Q6M8D0
58	545	18.8	552	2	Q25003
59	544	18.8	510	2	Q7UUK9
60	544	18.8	540	2	Q9MAW4
61	543	18.7	540	2	Q8LG88
62	540	18.6	462	2	Q9FGK1
63	536	18.5	548	2	Q9ZML8
64	533	18.4	626	2	Q8AQ77
65	528	18.2	587	2	Q8FU27
66	527	18.2	540	2	Q8FEH9
67	520.5	18.0	524	2	Q6NKS9
68	517.5	17.9	474	2	Q65MM0
69	496.5	17.1	450	2	Q75MH3
70	474	16.4	450	2	Q8FMZ5
71	463.5	16.0	302	2	Q68D44
72	459.5	15.9	461	1	Y608_HAEIN
73	447	15.4	470	2	Q7N822
74	446.5	15.4	474	2	Q8GF70
75	444.5	15.3	456	2	Q7MLG6
76	444.5	15.3	456	2	Q8D908
77	442.5	15.3	463	2	Q9CLO0
78	442	15.3	446	2	Q66449
79	436	15.1	461	1	Y608_HAEIN
80	433	14.9	462	2	Q9KNE0
81	428.5	14.8	459	2	Q66615
82	428.5	14.8	459	2	Q8C2T3
83	427.5	14.8	456	2	Q6LQW4
84	425	14.7	456	2	Q87Q94
85	425	14.7	471	2	Q9JUV43
86	425	14.7	471	2	Q9K032
87	420.5	14.5	456	2	Q8ZHX0
88	418	14.4	464	2	Q8VXT1
89	418	14.4	472	2	Q8DAD2
90	416	14.4	464	2	Q984F3
91	415.5	14.3	463	2	Q65SS5
92	415	14.3	464	2	Q8F5L4
93	415	14.3	472	2	Q7MJT9
94	414	14.3	464	2	Q9AED9
95	414	14.3	464	2	Q72QJ5
96	399	13.8	464	2	Q9ZGU3
97	396	13.7	471	2	Q6LSU9
98	396	13.7	487	2	Q9KSE3
99	388.5	13.4	451	2	Q7M8Y1
100	386	13.3	491	2	Q847W4

ALIGNMENTS

```

RESULT 1
ID_2 DROME STANDARD; PRT; 562 AA.
AC Q9VQDQ; Q95783;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE I'm not dead yet protein 2.
GN Name=Indy-2;
OS Drosophila melanogaster (Fruit fly);
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fobel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hastin N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Houston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laio K., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RT Science 287:2185-2195(2000).
RN [2]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.-J., Drysdale R.A.,
RA Harris N.L., Richter J.S., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Head, and Testis;

```

```

RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource."
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RN [4]
RP IDENTIFICATION.
RX PubMed=1118146; DOI=10.1126/science.290.5499.2137;
RA Rogina B., Reenan R.A., Nilsen S.P., Helfand S.L.;
RT "Extended life-span conferred by cotransporter gene mutations in
RL Drosophila."
RL Science 290:2137-2140(2000).
CC -!- FUNCTION: Cation-independent electroneutral transporter (not
CC associated with membrane depolarization) of a variety of
CC tricarboxylic and dicarboxylic acid-cycle intermediates. There is
CC also small, but detectable, transport of monocarboxylics.
CC Transport is through the epithelium of the gut and across the
CC plasma membranes of organs involved in intermediary metabolism and
CC storage (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: Belongs to the SLC13A transporter family. NADC
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AS003728; AAN13819.1; -
CC EMBL; AT060287; AAL25326.1; -
CC EMBL; AT060329; AAN71084.1; -
CC FlyBase; Fgn0011603; Indy-2.
CC GO; GO:0005887; C:integral to plasma membrane; ISS.
CC GO; GO:0015137; F:cyturate transporter activity; ISS.
CC GO; GO:0050833; F:pyruvate transporter activity; ISS.
CC GO; GO:0015141; F:succinate transporter activity; ISS.
CC GO; GO:0015746; P:citrate transport; ISS.
CC GO; GO:0008340; P:determination of adult life span; ISS.
CC GO; GO:0006848; P:pyruvate transport; ISS.
CC GO; GO:0015744; P:succinate transport; ISS.
CC InterPro; IPR001898; Na/sul symport.
CC Pfam; PF00939; Na sulph sym; 1.
CC PROSITE; PS01271; NA_SULFATE; FALSE_NEG.
CC Transmembrane; Transp.
CC TRANSMEM 24 44 Potential.
CC TRANSMEM 64 84 Potential.
CC TRANSMEM 93 113 Potential.
CC TRANSMEM 135 155 Potential.
CC TRANSMEM 200 220 Potential.
CC TRANSMEM 249 269 Potential.
CC TRANSMEM 308 328 Potential.
CC TRANSMEM 350 370 Potential.
CC TRANSMEM 393 413 Potential.
CC TRANSMEM 430 450 Potential.
CC TRANSMEM 478 498 Potential.
CC TRANSMEM 515 535 Potential.
CC SQUENCE 562 AA; 61934 MW; D9F04C05FAA73ED3 CRC64;

```

```

Query Match 100.0%; Score 2897; DB 1; Length 562;
Best Local Similarity 100.0%; Pred. No. 5.5e-187;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAEPGEORKEVLCGCCIFHWRGKASIIIPILITIPILYGFQTDMAEFKCLWLVITMALLW 60
Db 1 MAEPGEORKEVLCGCCIFHWRGKASIIIPILITIPILYGFQTDMAEFKCLWLVITMALLW 60
Qy 61 ITETLPPIVVTALPFLVFCPLLGVLNVAISIVCKQYFTDTIVVFLGGLIVALGIEYSNLHTRI 120
Db 61 ITETLPPIVVTALPFLVFCPLLGVLNVAISIVCKQYFTDTIVVFLGGLIVALGIEYSNLHTRI 120

```

QY 121 ALRVIRIVGSPRLFLVGLMSVSTFNGLWISNSAGTAMMCPVKAALVNELDNKNIPFVVM 180  
DB 121 ALRVIRIVGSPRLFLVGLMSVSTFNGLWISNSAGTAMMCPVKAALVNELDNKNIPFVVM 180  
QY 181 TOEEBVEGEPPHPSKITVAFVAGIAYASSIGGLTGLTGTLNLFVRCIYTERPPTSTV 240  
DB 181 TOEEBVEGEPPHPSKITVAFVAGIAYASSIGGLTGLTGTLNLFVRCIYTERPPTSTV 240  
QY 241 EITFANFMFYSIPLMVIVNVTLVIAFLITHMGLFRPNSTKGIABANTNKLKLMEDVLR 300  
DB 241 EITFANFMFYSIPLMVIVNVTLVIAFLITHMGLFRPNSTKGIABANTNKLKLMEDVLR 300  
QY 301 QRHIDLGPMSCHEIOAIAFAFMVLLITRKPGFVPGWSDLNKRKVGSGSLSFIVLLI 360  
DB 301 QRHIDLGPMSCHEIOAIAFAFMVLLITRKPGFVPGWSDLNKRKVGSGSLSFIVLLI 360  
QY 361 FALPTQYTFEKKYCCGKGPPTAQDAIDALSWEYVLRNIPWGLLFLGGGFPALAVASRETGL 420  
DB 361 FALPTQYTFEKKYCCGKGPPTAQDAIDALSWEYVLRNIPWGLLFLGGGFPALAVASRETGL 420  
QY 421 NIMISKAMQVLGLNIVVQSIITFVLNFFSAFNANVVANIVLPILCMSLALHPLI 480  
DB 421 NIMISKAMQVLGLNIVVQSIITFVLNFFSAFNANVVANIVLPILCMSLALHPLI 480  
QY 481 LTPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIYPTIIGISVALVNTNTWGL 540  
DB 481 LTPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIYPTIIGISVALVNTNTWGL 540  
QY 541 IIPPEKSPDPAWKEIKNQTKI 562  
DB 541 IIPPEKSPDPAWKEIKNQTKI 562

RESULT 2

Q811B6 PRELIMINARY; PRT; 505 AA.  
AC Q811B6  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Indy-2-PA.  
GN Name=Indy-2;  
OS Drosophila pseudoobscura (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7237;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Tucson 14011-0121.4;  
RA Bergman C.M., Pfeiffer B.D., Rincon-Limas D.E., Hoskins R.A., Park S.,  
RA Gnietke A., Mungall C.J., Wang A.M., Krommiller B., Pacleb J., Rubin G.M.,  
RA Stapleton M., Wan K., George R.A., de Jong P.J., Botas J., Rubin G.M.,  
RA Celniker S.E.;  
RT functional annotation of the Drosophila genome."  
RL Genome Biol. 3:research0086-research0086(2002).  
DR EMBL; AY199949; AAO01056.1; -;  
DR FlyBase; FBgn0064420; DpseIndy-2.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006814; P:sodium ion transport; IEA.  
DR InterPro; IPR001395; Aldo/ket\_red.  
DR InterPro; IPR001898; Na/sul\_symport.  
DR Pfam; PF00939; Na\_sulph\_sympt; 2.  
DR PROSITE; PS00663; ALDOXETO\_REDUCTASE\_3; UNKNOWN 1.  
SQ SEQUENCE 505 AA; 55608 MW; C1F8E3D37C579A40 CRC64;

Query Match 53.3%; Score 1543; DB 2; Length 505;  
Best Local Similarity 56.3%; Pred. No. 7.8e-96;  
Matches 285; Conservative 96; Mismatches 123; Indels 2; Gaps 2;

QY 50 LMLIVTMALLMITETPLIYVITALLPPLVFCPLGLLVNASIVCKQYQYTDITVIFLGLLVAL 109  
DB 1 MYLVNLMFALITEAIPLYLSLFPVFLPLFDILGSDQVCKLYFSDTVVMFVIGGLIAL 60  
QY 110 GIEYSLNLTALRALVIRIVGSPRLFLVGLMSVSTFNGLWISNSAGTAMMCPVKAALVNE 169  
DB 61 ALEYSNLHQRJAMKILLVGCSPRLHFLGVVVTFCISLWISNSAATAMMCPVKAALVNE 120  
QY 170 LDTNKFVPMYTOEEBVEGEPPHPSKITVAFVAGIAYASSIGGLTGLTGTLNLFVRC 229  
DB 121 LDSQNIPTVYKQEEBVEGEPPHPSKITVAFVAGIAYASSIGGLTGLTGTLNLFVRC 180  
QY 230 IYTERPPTSTVEITFANFMFYSIPLMVIVNVTLVIAFLITHMGLFRPNSTKGIABAN 289  
DB 181 LYDTRFPKSKTQIDFPIFMAIYPIVIVNVLLYFSLQVTHMGLFRGKSGTGLEVVRGT 240  
QY 290 TNRKLMEDVLRHIDLGPMSCHEIOAIAFAFMVLLITRKPGFVPGWSDLNKRKVGSG 349  
DB 241 EGQAVVKTIVIKARHQLGPMTCHEIQVTLFVLMVFLFTRKPGFVPGWDFLNAQKIGS 300  
QY 350 ASGLSFIVLILFALPTQYTFEKKYCCGKGPPTAQDAIDALSWEYVLRNIPWGLLFLGGGF 409  
DB 301 GPPVWLFWMLFALPTQYTFEKKYCCGKGPPTAQDAIDALSWEYVLRNIPWGLLFLGGGF 359  
QY 410 ALAVASRETGLNIMISKAMQVLGLNIVVQSIITFVLNFFSAFNANVVANIVLPILC 469  
DB 360 ALAEGSKVSGMAKMLGSLKFAKMPPIVVEGMCILIGLPCFATFSSNAICNIPISF 419  
QY 470 MSLALEHPLILTPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIYPTIIGIS 529  
DB 420 MALAIKHPKLTLTPPSLAISMAVHLPVSTPPNAIISGYAGIKTKYLALAGILPTIWA 479  
QY 530 VALVNTNTWGLIIPPEKSPDPAWKE 555  
DB 480 VLMLNSQ-YGNIIPESTKFPDWA 504

RESULT 3

Q86B89 PRELIMINARY; PRT; 504 AA.  
AC Q86B89  
DT 01-JUN-2003 (TREMBlrel. 24, Created)  
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE CG32921-PB (CG32921-pd).  
GN ORFNames=CG32921;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer P.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Aehburner M., Henderson S.N.,  
RA Suton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.B.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Frannkoch C., Baldwin D.,  
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.B., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,









Query Match		30.7%; Score 890; DB 2; Length 605;
Best Local Similarity		34.9%; Pred. No. 8.9e-52;
Matches		204; Conservative 105; Mismatches 207; Indels 68; Gaps 12;
QY	25	SIILPILITPLIYGFOTDMAEFKCLMLVITMALLWITETLPIYVVTALFPLVFCPLGLGLV 84
DB	17	SILFILLPLPLL---VPTKEAKCAYSIILMALLWCTETLPLAVTAFILFIFFPMGIM 73
QY	85	NASIVCKQYFTDITVIFGLGLIHALGHEYSNLHTRIALRVIRVGGSPRFLFGLMSVST 144
DB	74	DASEVSEYKLDITNIFLIGGLLVAIAVEHNLHTRIALRVLLIIGVRFAPFLGLFMMVTA 133
QY	145	FMGLWINSAGTAMCPIVVALNVEL-----DT----- 172
DB	134	FLSMWISNTATTAMVIAHVALBQLHKGPBEKDTENGHNVSFELQEPHGNKPEPSSL 193
QY	173	-----NKIFFVYMTQEEPEVEEG-----EPHPSKITVAFYAGIYASSIGGLGTLI 219
DB	194	REKENSVPVPTSMPEYKEKEEKEKEKEH-FKLSQMSLVCVYAASIGGIATLT 252
QY	220	GTGTLNLFVGRGTYRPTSTVEITFANFMFYSIPLMVIVNLTIVIAFL---ITHMGL-F 275
DB	253	GTTLNVLQGMNSLFPKNSVNFASFSGFAPFTWVL-----LILLSWILQLFLGFNF 308
QY	276	RPNSTKGTIAEANTNRKLMEDVLRQRHIDLGPMSCHIEOMIAFAPMIVLLITRKPGFV 335
DB	309	RKNFNCER---KSOEKERAAQYVLOTBKKLGPMSFAEIAVTFLFLLVLTFRPGFF 365
QY	336	PGWSDLI-----NRKVVGSAGSLFVLLIFALPTQYTFPKYCGKGPPTAQDAIDLSW 390
DB	366	PGWGNVAFSNEBGSMSVSDGTVAFFIILFIVPSLVPFSQQGSKF--KAPPALLDW 423
QY	391	EYVLNTPWGLLFLGGGFALAVASRETGLNIMISKAMQVLIGLPNIVQSIITFVLNPF 450
DB	424	NTVKNKMPNWLFLGGGFALAKGSEVSGLSMWLGNKLTPLQSIPIPAIFILCLVATF 483
QY	451	SAFNAVNVANIVLPILCENSLALELHPLITLTPACLGISGVYPLVSTTPNNAIVTQYAH 510
DB	484	TECASNAVTTILFPLILASMAQAICNPLYMVECTLSASLAPMLVATPFPNAIVSYGQ 543
QY	511	IKTKYFACCGIVPTIIGISVALVNTNTWGLIIFPESKSPFDWAK 554
DB	544	LKVIDMAKTGFLNIIGVLTITLAINTWSYPIF-QLDQDFTWQAQ 586
RESULT 8		
ID	Q6AZR9	PRELIMINARY; PRT; 586 AA.
AC	Q6AZR9;	
DT	25-OCT-2004	(TrEMBLrel. 28, Created)
DT	25-OCT-2004	(TrEMBLrel. 28, Last sequence update)
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)
DE	Slc13a2-prov protein.	
GN	Name=slc13a2-prov;	
OS	Xenopus laevis (African clawed frog).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;	
OC	Xenopodinae; Xenopus.	
OX	NCBI_TaxID=8355;	
RC	SEQUENCE FROM N.A.	
RC	TISSUE=Kidney;	
RX	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;	
RA	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.	
RT	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative."	
RL	Dev. Dyn. 225:384-391 (2002).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Kidney;	
RX	PubMed=12477932; DOI=10.1073/pnas.242603899;	
Query Match		30.4%; Score 880.5; DB 2; Length 586;
Best Local Similarity		35.2%; Pred. No. 3.8e-51;
Matches		201; Conservative 101; Mismatches 210; Indels 59; Gaps 9;
QY	30	LITPLIYGFOTDMAEFKCLMLVITMALLWITETLPIYVVTALFPLVFCPLGLGNASIV 89
DB	24	LPLPLLI---QT---KEASCAVILMAYVTCVTEVIPLAVTALLPVLFPFMGILESKEV 78
QY	90	CKQYFTDITVIFGLGLIHALGHEYSNLHTRIALRVIRVGGSPRRLFVGLMSVSTPMGLW 149
DB	79	CMQYKLDITNIFLIGGLIHALGHEYSNLHTRIALRVIRVGGSPRRLFVGLMSVSTPMGLW 138
QY	150	ISNSAGTAMCPIVVALN-----ELDTNKIFPVYMTQ----- 182
DB	139	ISNTATTAMVPIVQAVLSQLHTAEEDPSMLSELEGQTNPALETKNAIPMLQVTVSN 198
QY	183	-----EERPVEEGEPHPSPKIVAFYAGIYASSIGGLGTIGTGNLVFRGIYTERFP 236
DB	199	GHPVEPVEEPDEKDKHTSK---GMMLCVCAVSIIGGTATITGTGNLVKQGFQIFP 255
QY	237	TSTVEITFANFMFYSIPLMVIVNLTIVIAFLITHMGLFRPNSTKGTIIAEANTNRKLM 296
DB	256	NGMDILNFAWFGFAFPNNVIM---LCLSWFLQFSFGFNFKKTGWCGCATASEKRAAY 312
QY	297	DVLRQRHIDLGPMSCHIEOMIAFAPMIVLLITRKPGFVPGWSDLI-----NRKVVGSAG 352
DB	313	SVIREEYRKLGPISVAESSVFLFLILLVLLWFTDRDPGVVWATILFNKDNIETATDV 372
QY	353	LSFVILLIFALPTQYTFPKYCGKGPPTAQAD-----AILLSWEYVLNRNIPWGLL 402
DB	373	AVFVAELFILPATKPKFOCCAKNSFDEDEEBEQDFTFFSAPLITWVQKQMPWSIV 432
QY	403	FLIGGGFALAVASRETGLNIMISKAMQVLIGLPNIVQSIITFVLNFFAFSFAFNNAVANI 462
DB	433	LLUGGGFALAKGSDASGLSHWLQGMTPHLSIPPPWAIILSLMIAVFTFECASNATATL 492
QY	463	VLPLCENSLALELHPLITLTPACLGISGVYPLVSTPNAIVTQYAHITKTKYFACCGIV 522
DB	493	FLPILASMSRSIEVNPPLYIMIPCTLTSTSPAFMLPVATPFPNAIVFSGHLRVSMDVMTGIV 552

RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Whiting J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.C., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grumwood J., Schmutz J., Myers J.M., Butterfield Y.S.,
RA	Krzywinski M.I., Skalek U., Smailus D.E., Scherch A., Schein J.E.,
RA	Jones S.J., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences."
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Kidney;
RC	Klein S., Gerhard D.S.;
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BC077435; AAH77435.1; -
DR	GO; GO:0016020; C:membrane; IEA.
DR	GO; GO:0005215; F:transporter activity; IEA.
DR	GO; GO:0006814; P:sodium ion transport; IEA.
DR	InterPro; IPR001898; Na/sul_symport.
DR	Pfam; PF00939; Na sulph sym; 1.
DR	PROSITE; PS01271; NA_SULFATE; 1.
SQ	SEQUENCE 586 AA; 64959 MW; 2E3646A1D81F31DB CRC64;

QY 523 PFIIGISVALVNTWGLIIPESKSPFDWA 553  
 DB 553 MNIIGICTTISINTWGRPMF-SLDTFPQWA 582

RESULT 9  
 Q86YTS PRELIMINARY; PRT; 568 AA.  
 AC Q86YTS;  
 DT 01-JUN-2003 (Tremblrel. 24, Created)  
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Na+-coupled citrate transporter protein.  
 GN Name=NaCT;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22334959; PubMed=12445824; DOI=10.1016/S0006-291X(02)02669-4;  
 RA Inoue K., Zhuang L., Ganapathy V.;  
 RT "Human Na+ -coupled citrate transporter: primary structure, genomic  
 organization, and transport function.";  
 RL Biochem. Biophys. Res. Commun. 299:465-471(2002).  
 DR EMBL; AY151833; AAN86530.1; -;  
 DR Genbank; HGNC:23089; SLC13A5.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006814; P:sodium ion transport; IEA.  
 DR InterPro; IPR001898; Na/sul symport.  
 DR Pfam; PF00939; Na sulph sym; 1.  
 DR PROSITE; PS01271; NA SULFATE; 1.  
 SQ SEQUENCE 568 AA; 63062 MW; B8995E5618DECCB CRC64;

Query Match 30.2%; Score 875; DB 2; Length 568;  
 Best Local Similarity 35.7%; Pred. No. 8.6e-51;  
 Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;

QY 27 IIPILITPILYIGFTDMAEP-KCLWLIVTMALLWITETPLIYVYALPFLVPCPLGLVN 85  
 DB 19 VTPLLLPVLI-----LMPAKFVRCAYIILMAIYCTEIVPLAVTSLMPVLLPFLQILD 74  
 QY 86 ASIVCKQYFTDTIVVFLGGLIYALGIEYSNLHTRIALRVIRVGGSPRLVGLMSVTF 145  
 DB 75 SRQVCVQWKDTNMLFLGLLIVAVAVERNLHKLIALRTLLVWGAKPARLMFGMGVTL 134  
 QY 146 MGLWISNAGTAMMCPIVKALVNELDN-----KIFPVYMTQBEET-VE 188  
 DB 135 LSMWISNTATTAMVPIVEAILQQMEATSAATEAGLELVKDGKAKELPGSQVIFEGPTLG 194  
 QY 189 EGEPPHPSKITVAFYAGIAYASSIGGLTGLTGTNLVFRGIYTERPTSTVEITFANFM 248  
 DB 195 QQEDQERKRLCKAMTLCICYAASIGGTATLTGTGNVLLGQNNELFPDSDKLVNFASF 254  
 QY 249 FYSIPLMIVNVTVIIAFLITHMGLRPN-SKTGKIIAEANTNRKLMEDVLRORHIDL 307  
 DB 255 AFAPNMLV-----MLLPAWMLQFYVNFNFKKSWGCGLESKQNEKAAKVLQBEYRKL 310  
 QY 308 PMSCHEIOMAFAMIVLLITRKPGFVPGWSDLI-----NRKVGSASGLSFVILLIPAL 363  
 DB 311 PLSFAEINVLCIFLLVILFSDRDPGMPGLTVAVVEGETKYVSDATVAIFVATLLFIV 370  
 QY 364 PTQYTFKYCC-----GKGPTAQDAIILSWEYVLRNIPNGLLFLGGGFPALAVASRET 418  
 DB 371 PSQPKFNFRSQTBEERTFPY---PPLLDWKVTQEKVPMGIVLLGGGFPALAKGSEAS 427  
 QY 419 GLNIMISAMQVILGLPNI VQSITFVLNFFSFAFNANVAVNI VLPILCMSLALHP 478  
 DB 428 GLSVWGMQKQEPHVAVPAATILLSLVAVFECTSNVATTTFLPFIASMSRSIGLNP 487  
 QY 479 LILTLPACLGISVMYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTTIIGISVALVNTW 538

DB 488 LYIMLPCTLSASFAPMLFPVATPPNAIVTYGHLKVAQDMVKTGVMNIIIGVFCVFLAVNTW 547  
 QY 539 GLIIFPESKSPFDWA 553  
 DB 548 GRAIF-DLDHFPDWA 561

RESULT 10  
 Q6ZMG1 PRELIMINARY; PRT; 568 AA.  
 AC Q6ZMG1;  
 DT 05-JUL-2004 (Tremblrel. 27, Created)  
 DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)  
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
 DE Hypothetical protein FLJ23946.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kawakami T., Noguuchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,  
 RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,  
 RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK172785; BAD18766.1; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005215; P:transporter activity; IEA.  
 DR GO; GO:0006814; P:sodium ion transport; IEA.  
 DR InterPro; IPR001898; Na/sul symport.  
 DR Pfam; PF00939; Na sulph sym; 1.  
 DR PROSITE; PS01271; NA SULFATE; 1.  
 SQ SEQUENCE 568 AA; 63060 MW; -2535BEF045C698DB CRC64;

Query Match 30.1%; Score 871; DB 2; Length 568;  
 Best Local Similarity 35.7%; Pred. No. 1.6e-50;  
 Matches 198; Conservative 98; Mismatches 219; Indels 40; Gaps 10;

QY 27 IIPILITPILYIGFTDMAEP-KCLWLIVTMALLWITETPLIYVYALPFLVPCPLGLVN 85  
 DB 19 VTPLLLPVLI-----LMPAKFVRCAYIILMAIYCTEIVPLAVTSLMPVLLPFLQILD 74  
 QY 86 ASIVCKQYFTDTIVVFLGGLIYALGIEYSNLHTRIALRVIRVGGSPRLVGLMSVTF 145  
 DB 75 SRQVCVQWKDTNMLFLGLLIVAVAVERNLHKLIALRTLLVWGAKPARLMFGMGVTL 134  
 QY 146 MGLWISNAGTAMMCPIVKALVNELDN-----KIFPVYMTQBEET-VE 188  
 DB 135 LSMWISNTATTAMVPIVEAILQQMEATSAATEAGLELVKDGKAKELPGSQVIFEGPTLG 194  
 QY 189 EGEPPHPSKITVAFYAGIAYASSIGGLTGLTGTNLVFRGIYTERPTSTVEITFANFM 248  
 DB 195 QQEDQERKRLCKAMTLCICYAASIGGTATLTGTGNVLLGQNNELFPDSDKLVNFASF 254  
 QY 249 FYSIPLMIVNVTVIIAFLITHMGLRPN-SKTGKIIAEANTNRKLMEDVLRORHIDL 307  
 DB 255 AFAPNMLV-----MLLPAWMLQFYVNFNFKKSWGCGLESKQNEKAAKVLQBEYRKL 310  
 QY 308 PMSCHEIOMAFAMIVLLITRKPGFVPGWSDLI-----NRKVGSASGLSFVILLIPAL 363  
 DB 311 PLSFAEINVLCIFLLVILFSDRDPGMPGLTVAVVEGETKYVSDATVAIFVATLLFIV 370  
 QY 364 PTQYTFKYCC-----GKGPTAQDAIILSWEYVLRNIPNGLLFLGGGFPALAVASRET 418  
 DB 371 PSQPKFNFRSQTBEERTFPY---PPLLDWKVTQEKVPMGIVLLGGGFPALAKGSEAS 427  
 QY 419 GLNIMISAMQVILGLPNI VQSITFVLNFFSFAFNANVAVNI VLPILCMSLALHP 478  
 DB 428 GLSVWGMQKQEPHVAVPAATILLSLVAVFECTSNVATTTFLPFIASMSRSIGLNP 487  
 QY 479 LILTLPACLGISVMYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTTIIGISVALVNTW 538



Matches 211; Conservative 100; Mismatches 205; Indels 61; Gaps 16;

Qy 20 WRGKASIII---PLITLPILIIYGFQTDMAEFKCLWLIVTMALLWITETLPIVYTALPFLV 76  
  ::: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:  
Db 9 WAYRSYLIVLCIFILPLI---VQTKEAYCAYSIIIMALLWCTEALPLAVTALPFII 65

Qy 77 FCPILGLUNASIYCKQYFTDTIVVFLLGLIIVALGIYSNLUHTRIALRVIRVGSPRRLF 136  
  ::: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:  
Db 66 LFFPLMGIMEASKVCLEFYFKDNTLTFVGGLMVAIVAEHNHLKRITALGVLLIIIGVRPALLL 125

Qy 137 VGLMSYSTFNGLIWSNSAGTAMCPVTAKLVNEL-----DTNKIPPVVMTOEEEPVEE 189  
  ::: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:  
Db 126 LGFMVLTAFLSMWISNTATTMMPLTCGYAVELOQSQDKQVEEGSNPSFELQEASPOKE 185

Qy 190 -----GEP-PPHSKITVAFYAG-----IAYASSIGGLTGTLIGTNLNVFRGIY 231  
  ::: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:  
Db 186 ETKLDNGQAVSVSESPRAQKTKEHRFSQSLSCICVSASIGGTATLTGTTPNLVLOQOV 245

Qy 232 TERFPPTSTVEITANFMFYSIPLMVIWNVTLVIAEL---ITHMGL-FRNSKTYGKIIE 287  
  ::: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:  
Db 245 NSIFPENSNVNFNAPSGFGFAFTPMVI-----LLLMLQLVQLFVGNFRKNFGFG----E 297

Qy 288 ANTNRKIME-DVLRQRHIDIGPMSCHEIOAIAFAPFWILLIITRKPGFVPGWSD--LINR 344  
  ::: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:  
Db 298 GEERQAAPQVVIKTHRUJLGPMSPAETVFLVLLVLMFTREPFGFPGWGDTAFANK 357

Qy 345 K---VVGSAGLSFIIVLLIPALTQTVTFKYCG-----KGPTAAQADAILSWHEYVLRN 396  
  ::: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:  
Db 358 KGQSMVSDGTVAIFISLIWPIIPSkip-----GLTEDPKKPGKLKAPPAILTWKTVDNK 411

Qy 397 IPWGLLFLGGGFPALAVASRETGLNIMI SKAQVLI GLNPVVQSITFVLANPFSAFNAN 456  
  ::: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:  
Db 412 MPNNIILILGGGFPALAKGEESGLSKWLGDKLTPLOHPVP SATVLILLSLVAIFTECTSN 471

Qy 457 VVNAVNI VLPTLCBMSLALELHPILITLPACLGISMVVFVLPSTPPNAIVTQYAHIKTKYF 516  
  ::: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:  
Db 472 VATTLFLPLASMAQAI CLHPLVYMLPCTLAASLAFMLP VATPPNAIVSFPGGLKVSDM 531

Qy 517 ACGIIVPTTIIGISVALVNTNWGLIIPFSKSPDDWA 553

Db 532 ARAGFLNIIGVLTITLSINSWSIPIF-KLDTPTTWA 567

RESULT 13

8QCJ44 PRELIMINARY; PRT; 572 AA.

ID Q8CJ44 Q8CJ44; Q8CJ44;  
DC 01-MAR-2003 (TEMBLrel. 23, Created)  
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
DE Sodium-coupled citrate transporter.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxId=10116;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
RP MEDLINE=22267306; PubMed=12177002; DOI=10.1074/jbc.M207072200;  
RX Inoue K., Zhuang L., Maddox D.M., Smith S.B., Ganapathy V.;  
RT "Structure, function, and expression pattern of a novel sodium-coupled  
RT citrate transporter (NaCT) cloned from mammalian brain.";  
RL J. Biol. Chem. 277:39469-39476(2002).  
DR EMBL; AF522186; AAC52081.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006814; P:sodium ion transport; IEA.  
DR InterPro; IPR001898; Na/sulc symport.  
DR Pfam; PF00939; Na\_sulph\_sympt. 1.  
DR TIGRFAMS; TIGR00785; dase; 1.  
DR PROSITE; PS01271; NA\_SULFATE; 1.  
SQ SEQUENCE 572 AA; 63878 MW; 9CD35A72FC50D374 CRC64;



27	QY	IIPLITLPIILYIGFOTDMAEKPCKLWLIWVWALLMI	TETPLIYVYALPPLVFCPLLGVNA	86
		::: :::	::: :::	
	Db	19 LVPLFLLPLV---	VPTKEASCGFVIVWALFCTEALPLAVTALFVLLVFPFMGIMDS	75
		::: :::	::: :::	
87	QY	SIVCQYQYFDTIVVFLGGLIYVALGLEYISNLHTRIALREIVIRIVGSGPRRLFVGLMSYSTVM	146	
		::: :::	::: :::	
76	Db	TAVCSQYLKDNMLPIGGLAVAI	SVEKWNLHKRIALRAVLII	135
		::: :::	::: :::	
147	QY	GLWTSNSAGTAMMCPVIVKALVNELDTK-		177
		::: :::	::: :::	
136	Db	SNWISNTATTAMTPIAQAVALQQLHTSEER	VDKPSBETKKNVNGMONSMYESVLPNGKMI	195
		::: :::	::: :::	
178	QY	---VYMTQEE--	EPVEEGEPHPHS	198
		::: :::	::: :::	
196	Db	LAVENTYATVNEGFEOMODSTKDPEFSKLEKESIG	PIVIELEDEKQKEEKQKRLKHLKXI	255
		::: :::	::: :::	
199	QY	TVAFVAGIAYAVASSIGGLTIGCTNLVFRGIYTEREPTSTVEIT	FANFMFYSLPML-VI	257
		::: :::	::: :::	
256	Db	CKGMTLSVCYASIGGIATLTGTTPLNVKGMDELFPENN	VINVFASWFGPAFTMFLV	315
		::: :::	::: :::	
258	QY	VNVTLVLIIFALITHMGL--	FRPNSKTGKIIAEANTNRKLM-	312
		::: :::	::: :::	
316	Db	LALSMLWLOQFNF--	LGFNFKNFQCG----	368
		::: :::	::: :::	

Search completed: June 30, 2005, 08:58:43  
Job time : 77.6279 secs

Search completed: June 30, 2005, 08:58:43  
Job time : 77.6279 secs

**THIS PAGE IS BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:14 ; Search time 82.8208 Seconds  
(without alignments)  
2741.199 Million cell updates/sec

Title: US-10-017-479a-4

Perfect score: 3058

Sequence: 1 MATCHPALWAYRFXYLIVLCL.....HSNTSQCLLNPSNVTVPGL 587

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A\_Geneseq\_16Dec04.\*

1: Geneseqp1980a.\*

2: Geneseqp1990a.\*

3: Geneseqp2000a.\*

4: Geneseqp2001a.\*

5: Geneseqp2002a.\*

6: Geneseqp2003a.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004a.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3058	100.0	587	7	Aae38766 Rat 69624
2	2392	78.2	592	6	Abb82950 Human SLC
3	2392	78.2	592	6	Abb82951 Human SLC
4	2392	78.2	592	6	Abb82951 Human SLC
5	2392	78.2	592	6	Abb82951 Human SLC
6	2392	78.2	592	6	Abb82951 Human SLC
7	2073.5	67.8	519	8	Adh22543 Human tra
8	2064.5	67.5	519	8	Adh22543 Human tra
9	1971.5	64.5	619	8	Adg16995 African c
10	1829	59.8	581	8	Adp64799 Zebrafish
11	1628	53.2	616	5	Abu65064 Human NOV
12	1628	53.2	616	5	Abu65063 Human NOV
13	1628	53.2	616	7	Adk51048 Human NOV
14	1628	53.2	616	8	Adh42447 Novel hum
15	1628	53.2	616	8	Adn61777 Human nov
16	1628	53.2	616	8	Adn61779 Human nov
17	1624	53.1	568	8	Abf57023 Human NOV
18	1624	53.1	568	7	Aae38764 Human 696
19	1624	53.1	568	7	Adk51052 Human NOV
20	1624	53.1	568	8	Adh42441 Novel hum
21	1624	53.1	568	8	Adp64793 Human Na+
22	1624	53.1	568	8	Adh42443 Novel hum
23	1622	53.0	568	5	Abu65062 Human NOV
24	1622	53.0	568	7	Adk51050 Human NOV
25	1622	53.0	568	8	Adh42445 Novel hum

26	1622	53.0	568	8	ADN61775	Adn61775 Human nov
27	1621	53.0	568	7	ADE07994	Novel pro
28	1613	52.7	568	6	ABG75835	Transport
29	1610	52.6	568	6	ABR40097	Human sod
30	1608	52.6	568	5	AAU79946	Human tra
31	1608	52.6	568	8	ADG16993	Human tra
32	1562.5	51.1	531	6	AGG75826	Transport
33	1559.5	51.0	539	5	AAE21181	Human TR1
34	1538.5	50.3	551	8	ADRO9550	Human pro
35	1536.5	50.2	572	6	ABR57024	Mouse TCH
36	1536.5	50.2	572	8	ADP64797	Mouse Na+
37	1504.5	49.2	572	8	ADP64791	Rat INDY
38	1408	46.0	522	5	ABU65065	Human NOV
39	1408	46.0	522	7	ADK51054	Human NOV
40	1408	46.0	522	8	ADH42449	Novel hum
41	1408	46.0	522	8	ADN61781	Human nov
42	1355	44.3	602	6	ABR82952	Human SLC
43	1355	44.3	602	6	ABR40101	Human ren
44	1355	44.3	602	6	ABG75837	Transport
45	1337	43.7	516	5	ABU65066	Human NOV
46	1337	43.7	516	7	ADK51056	Human NOV
47	1337	43.7	516	8	ADH42451	Novel hum
48	1325	43.3	595	5	AAE22910	Human tra
49	1325	43.3	595	5	AAO21807	Lung-spec
50	1325	43.3	595	6	ABR82949	Human SLC
51	1325	43.3	595	7	ADI21045	Novel hum
52	1325	43.3	595	8	ADO78130	Human SLC
53	1323.5	43.3	477	8	ADN61783	Human nov
54	1259	41.2	627	3	AAAB36161	Novel hum
55	1257	41.1	627	3	AAAB3625	Human sec
56	1257	41.1	627	3	AAAB36158	Novel hum
57	1256	41.1	627	3	AAAB42213	Human ORF
58	1256	41.1	627	5	ABR97450	Novel hum
59	1255.5	41.1	626	3	AAAB36167	Novel hum
60	1254	41.0	627	3	AAAB36162	Novel hum
61	1253.5	41.0	626	3	AAAB36164	Novel hum
62	1252	40.9	627	3	AAAB36159	Novel hum
63	1250.5	40.9	626	3	AAAB36168	Novel hum
64	1248.5	40.8	626	3	AAAB36165	Novel hum
65	1199.5	39.2	533	8	ABM84303	Human dia
66	1186	38.8	557	8	ADP29692	Human sec
67	1184.5	38.7	520	7	ADMO5177	Human pro
68	1168.5	38.2	552	6	AAO31005	Human tra
69	1097	35.9	581	3	AAAB36163	Novel hum
70	1095	35.8	581	3	AAAB36160	Novel hum
71	1093.5	35.8	580	3	AAAB36169	Novel hum
72	1091.5	35.7	580	3	AAAB36166	Novel hum
73	1042.5	34.1	572	4	ABR60315	Drosophil
74	1042.5	34.1	572	4	ABR60315	Drosophil
75	1042.5	34.1	572	5	ABR79611	Drosophil
76	1042.5	34.1	572	8	ADP64789	Drosophil
77	1039	34.0	382	7	ADP60941	Novel pro
78	1034.5	33.8	572	6	ABR40099	dindya s
79	1014.5	33.2	407	5	AAU91115	Human sec
80	1014.5	33.2	407	5	ABR65198	Human alb
81	1014.5	33.2	407	8	ADL78465	Albumin f
82	997	32.6	377	5	AAU91090	Human sec
83	997	32.6	377	5	ABR65200	Human alb
84	997	32.6	377	8	ADL78467	Albumin f
85	887.5	29.0	551	8	ADP64795	Nematode
86	849.5	27.8	585	7	AAE38765	Human 696
87	835.5	27.3	432	6	ABM71452	Staphyloc
88	831.5	27.2	466	3	AAE38765	Human 696
89	831.5	27.2	540	3	AAE38765	Human 696
90	831.5	27.2	558	3	AAE38765	Human 696
91	829.5	27.1	466	3	AAE38765	Human 696
92	829.5	27.1	540	3	AAE38765	Human 696
93	829.5	27.1	557	3	AAE38765	Human 696
94	799	26.1	543	6	ADA48682	Rice prot
95	799	26.1	543	7	ADC08241	Rice prot
96	787	25.7	908	4	ABR61068	Drosophil
97	781	25.5	552	2	AAW98815	H. pylori
98	776	25.4	527	4	AAW98815	H. pylori

99 773 25.3 510 4 AAG90000 Aag90000 C glutami  
100 706.5 23.1 304 5 ABB89646 ABB89646 Human pol

ALIGNMENTS

RESULT 1  
AAE38766  
ID AAE38766 standard; protein; 587 AA.  
XX  
AC AAE38766;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Rat 69624 protein sodium sulphate symporter domain.  
XX  
KW Rat; 69624; transporter protein; neurological disorder; therapy;  
KW atherosclerosis; cardiac hypertrophy; ischaemia reperfusion injury;  
KW metabolic disorder; haematopoietic neoplastic disorder; leukaemia;  
KW arthritis; multiple sclerosis; encephalomyelitis; myasthenia gravis;  
KW carcinoma; cell proliferation; autoimmune disorder; diabetes mellitus;  
KW renal disorder; colon; hepatic disorder; hypocalcaemia; calcium stone;  
KW mental retardation; Canavan disease; Differentiative disorder; sarcoma;  
KW systemic lupus erythematosus; cardiovascular disorder; arteriosclerosis;  
KW atrial fibrillation; forensic identification; pain.  
XX

OS Rattus norvegicus.  
XX  
XX US2002193582-A1.  
XX  
XX 19-DEC-2002.  
XX  
XX 17-JUN-2002; 2002US-00173519.  
XX  
XX 18-JUN-2001; 2001US-0298970P.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Curtis RAJ;  
XX  
XX WPI; 2003-644620/61.  
XX

Novel 69624 polypeptide, a human transporter family member, useful for treating disorders e.g. hypocalcaemia, such as hypocalcaemia, formation of calcium stones, mental retardation abnormal body sulfate homeostasis.  
XX  
PS Disclosure; Page 49-50; Opp; English.  
XX

The invention relates to 69624 polypeptide, a human transporter family member and its corresponding nucleic acid. 69624 protein is useful for developing novel diagnostic and therapeutic agents for 69624-mediated or related disorders. 69624 protein act as therapeutic or diagnostic agents for renal, neurological, colon or hepatic disorders. It act as diagnostic targets and therapeutic agents for treating disorders such as hypocalcaemia, formation of calcium stones, mental retardation (Canavan disease) or abnormal body sulphate homeostasis. 69624 protein may act as diagnostic targets and therapeutic agents for controlling cellular proliferative and/or differentiative disorders such as carcinoma, sarcoma, metastatic disorder or haematopoietic neoplastic disorders e.g., leukaemia, immune disorders such as autoimmune disorders (diabetes mellitus, arthritis), multiple sclerosis, encephalomyelitis, myasthenia gravis, systemic lupus erythematosus, cardiovascular disorders such as arteriosclerosis, atherosclerosis, ischaemia reperfusion injury, cardiac hypertrophy, atrial fibrillation etc; and disorders involving abnormal or excessive pain. 69624 sequence is useful as pharmacodynamic marker and is also used in forensic identification of a biological sample. The present sequence is rat sodium-dependent dicarboxylate/sulphate transporter (NaDC -1) protein

XX Sequence 587 AA;

Query Match 100.0%; Score 3058; DB 7; Length 587;

Best Local Similarity 100.0%; Pred. No. 7.7e-300;  
Matches 587; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATCHPALWAYRFLIVLCPIELPLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60  
DB 1 MATCHPALWAYRFLIVLCPIELPLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60  
QY 61 LFPVILFPLMGIMDASEVCIEYFKDTNIFLVGGLMVAIAVEHNLHKLALQVLLIIGVR 120  
DB 61 LFPVILFPLMGIMDASEVCIEYFKDTNIFLVGGLMVAIAVEHNLHKLALQVLLIIGVR 120  
QY 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLQSGKQVGGNNNPTFELQEE 180  
DB 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLQSGKQVGGNNNPTFELQEE 180  
QY 181 CPQKEVTKLDNGQPVASPEPRTKQEHHRFSQGLSLCICYSASIGGIATLTGTTPNLV 240  
DB 181 CPQKEVTKLDNGQPVASPEPRTKQEHHRFSQGLSLCICYSASIGGIATLTGTTPNLV 240  
QY 241 LQGVNSLFPQNGVNVNFASFQFAPPTMIILLALLMQLVFLGVNFRKNFGFGE 300  
DB 241 LQGVNSLFPQNGVNVNFASFQFAPPTMIILLALLMQLVFLGVNFRKNFGFGE 300  
QY 301 ERKQAAFOVIKQYRLLGPMSPFAEKTVTVLVLLVVLVFTREPPFPFGMDTVFANEKGQ 360  
DB 301 ERKQAAFOVIKQYRLLGPMSPFAEKTVTVLVLLVVLVFTREPPFPFGMDTVFANEKGQ 360  
QY 361 SMASDGTVAIFISLVNFIIPSKIPLGMQDPKPKGLKAPPAITLTKVNDKMPNIVILL 420  
DB 361 SMASDGTVAIFISLVNFIIPSKIPLGMQDPKPKGLKAPPAITLTKVNDKMPNIVILL 420  
QY 421 GGGFALAKGSEQSLSEWLGDKLTPLQHIPPSATAVILCLLIAIETECTSNVATTFLFLP 480  
DB 421 GGGFALAKGSEQSLSEWLGDKLTPLQHIPPSATAVILCLLIAIETECTSNVATTFLFLP 480  
QY 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPNAIVFSGGLKVSMDARAGFLNNI 540  
DB 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPNAIVFSGGLKVSMDARAGFLNNI 540  
QY 541 IGVLAITLSINSWSPIFKLDTFPPSWAHSNTSOCLLNPSNSTVPGGL 587  
DB 541 IGVLAITLSINSWSPIFKLDTFPPSWAHSNTSOCLLNPSNSTVPGGL 587

RESULT 2

AB82950  
ID AB82950 standard; protein; 592 AA.

XX  
AC ABB82950;  
XX  
DT 14-APR-2003 (first entry)  
XX  
DE Human SLC13A related protein (GenBank Identifier No. GI#2499523).  
XX  
KW SLC13A; p53; sodium-sulfate cotransporter 2; cytosolic; cancer;  
KW transmembrane protein; human.  
XX  
OS Homo sapiens.  
XX  
XX WO200298468-A1.  
XX  
PD 12-DEC-2002.  
XX  
PF 03-JUN-2002; 2002WO-US017460.  
XX  
PR 05-JUN-2001; 2001US-0296076P.  
PR 10-OCT-2001; 2001US-0328605P.  
PR 15-FEB-2002; 2002US-0357253P.  
PR 01-MAR-2002; 2002US-0361196P.  
XX  
XX (EXEL-) EXELIXIS INC.

Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

PI Lioubin MN;  
DR WPI; 2003-167297/16.  
XX  
XX  
XX Identifying candidate p53 pathway modulating agent for diagnosing or  
PT treating cancer comprises detecting test agent-biased activity of an  
PT assay system comprising purified Sodium sulfate cotransporter 2 (SLC13A)  
PT polypeptide.  
XX  
XX  
XX Claim 13; Page 59-61; 69pp; English.  
XX  
XX The invention relates to identifying a candidate p53 pathway modulating  
CC agent. The method involves assaying purified Sodium-sulfate cotransporter  
CC 2 (SLC13A) polypeptide or nucleic acid or a functionally active fragment  
CC or derivative with a test agent. The methods are useful for identifying a  
CC candidate p53 pathway modulating agent, modulating a p53 pathway of a  
CC cell, or a mammalian cell and for diagnosing a disease in a patient. The  
CC disease is breast, colon, lung or ovary cancer having greater than 25%  
CC expression level. The method is useful for manufacturing a medicament for  
CC diagnosing or treating breast, colon, lung or ovary cancer. Sequences  
CC AB82949-952 represent polypeptide sequences related to the human SLC13A  
CC protein  
XX  
XX Sequence 592 AA;  
Query Match 78.2%; Score 2392; DB 6; Length 592;  
Best Local Similarity 76.9%; Pred. No. 2e-232;  
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;  
Qy 1 MATCPALWAYRFLVILVCLPIFLPLPLIVQTEKAYCAYSIILMALLMCTEALPLAVTA 60  
Db 1 MATCQALWAYRSYLIVFFVPIILLPLILPLVPSKEAYCAYAIILMALPFWCTEALPLAVTA 60  
Qy 61 LFPILVPLMGIMDASEVCIEYFKDTNLFVGGMLVAIAVEHWNHKKRIALQVLLIIGVR 120  
Db 61 LFPILVPLMGIMDASEVAVELYKDSNLLFFGGLLVAIAVEHWNHKKRIALRVLLIIGVR 120  
Qy 121 PALLLGFMVLTAFLSMWISNTATTAMVPTIGHAVLEQLOGSK--KDYEGGNNPTFELQ 178  
Db 121 PAPLILGFMVLTAFLSMWISNTATSMVPTIAHAVLDQHSQASSNVBECSNPTFELQ 180  
Qy 179 BECPQKEVTKLDNGO--PV-SAPSEPTQKQEHFRFSQGLSLCICYSASIGGIATLTGT 235  
Db 181 EPSPOKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCHSLCVCYSASIGGIATLTGT 240  
Qy 236 TPNLVLOQVNSLFPQNGNVNFAWFSFAPPTWVILLALLAWMLQILFLGFNFRKNFGI 300  
Db 241 APNLVLOQVNSLFPQNGNVNFAWFSFAPPTWVILLALLAWMLQILFLGFNFRKNFGI 300  
Qy 296 GEGEERKQAAFOVITKQYRLLGPMSPAEKTVTVLVLLVLTFTREPFGPGMDTVFA 355  
Db 301 GEKQEQOQAAVCYVLTQEHRLGPMPTFAEKASILFVLVLLVLTFTREPFGFGNLAFF 360  
Qy 356 NEKGQSMASDGTVAIFISLVFIIPSKIPLGLMQDPKPKGLKAPPAIILTWVNDKMPWN 415  
Db 361 NAKGESWSDGTVAIFIGIINFILPSKFPGLTQDPENFGKLKAPGLGLDWTQKMPWN 420  
Qy 416 IVILGGGFALAKSEQGLSEWLDKLTPIQHTPPSATVILCLLIAIFTECTSNVATT 475  
Db 421 IVLLGGGYALAKSERSGLSEWLNKLTPLQSPVAPAIATILSLVATFTECTSNVATT 480  
Qy 476 TLFPIILASMAQAICLHPYVLMPCITLAASLAFMLPVATPNATVPNAIVFSPGGLKVSDBARAG 535  
Db 481 TIFPIILASMAQAICLHPYVLMPCITLATSALFMLPVATPNATVPNAIVFSGDLKVLDMARAG 540  
Qy 536 FLLNIIGVLTATLSINWSIPFKLDTFPPSAHNS--TSQCLINPNSNTPV 584  
Db 541 FLLNIIGVLTALAINSWGIPFLSLHSPSAQSNNTTAAQCLPSLANTTTP 590  
RESULT 3  
ABB82951  
ID ABB82951 standard; protein; 592 AA.

XX  
XX AC ABB82951;  
XX  
XX DT 14-APR-2003 (first entry)  
XX  
XX DE Human SLC13A related protein (GenBank Identifier No. GI#4506979).  
XX  
XX KW SLC13A; p53; sodium-sulfate cotransporter 2; cytostatic; cancer;  
XX transmembrane protein; human.  
XX  
XX OS Homo sapiens.  
XX  
XX FN WO200298468-A1.  
XX  
XX PD 12-DEC-2002.  
XX  
XX PF 03-JUN-2002; 2002WO-US017460.  
XX  
XX PR 05-JUN-2001; 2001US-0296076P.  
XX PR 10-OCT-2001; 2001US-0328605P.  
XX PR 15-FEB-2002; 2002US-0357253P.  
XX PR 01-MAR-2002; 2002US-0361196P.  
XX  
XX PA (EXEL-) EXELIXIS INC.  
XX  
XX PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;  
PI Lioubin MN;  
XX  
XX WIPI; 2003-167297/16.  
DR N-PSDB; ABZ24204.  
XX  
XX Identifying candidate p53 pathway modulating agent for diagnosing or  
PT treating cancer comprises detecting test agent-biased activity of an  
PT assay system comprising purified Sodium sulfate cotransporter 2 (SLC13A)  
PT polypeptide.  
XX  
XX PS Claim 13; Page 62-64; 69pp; English.  
XX  
XX The invention relates to identifying a candidate p53 pathway modulating  
CC agent. The method involves assaying purified Sodium-sulfate cotransporter  
CC 2 (SLC13A) polypeptide or nucleic acid or a functionally active fragment  
CC or derivative with a test agent. The methods are useful for identifying a  
CC candidate p53 pathway modulating agent, modulating a p53 pathway of a  
CC cell, or a mammalian cell and for diagnosing a disease in a patient. The  
CC disease is breast, colon, lung or ovary cancer having greater than 25%  
CC expression level. The method is useful for manufacturing a medicament for  
CC diagnosing or treating breast, colon, lung or ovary cancer. Sequences  
CC AB82949-952 represent polypeptide sequences related to the human SLC13A  
CC protein  
XX  
XX SQ Sequence 592 AA;  
Query Match 78.2%; Score 2392; DB 6; Length 592;  
Best Local Similarity 76.9%; Pred. No. 2e-232;  
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;  
Qy 1 MATCPALWAYRFLVILVCLPIFLPLPLIVQTEKAYCAYSIILMALLMCTEALPLAVTA 60  
Db 1 MATCQALWAYRSYLIVFFVPIILLPLILPLVPSKEAYCAYAIILMALPFWCTEALPLAVTA 60  
Qy 61 LFPILVPLMGIMDASEVCIEYFKDTNLFVGGMLVAIAVEHWNHKKRIALQVLLIIGVR 120  
Db 61 LFPILVPLMGIMDASEVAVELYKDSNLLFFGGLLVAIAVEHWNHKKRIALRVLLIIGVR 120  
Qy 121 PALLLGFMVLTAFLSMWISNTATTAMVPTIGHAVLEQLOGSK--KDYEGGNNPTFELQ 178  
Db 121 PAPLILGFMVLTAFLSMWISNTATSMVPTIAHAVLDQHSQASSNVBECSNPTFELQ 180  
Qy 179 BECPQKEVTKLDNGO--PV-SAPSEPTQKQEHFRFSQGLSLCICYSASIGGIATLTGT 235  
Db 181 EPSPOKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCHSLCVCYSASIGGIATLTGT 240  
Qy 236 TPNLVLOQVNSLFPQNGNVNFAWFSFAPPTWVILLALLAWMLQILFLGFNFRKNFGI 300  
Db 241 APNLVLOQVNSLFPQNGNVNFAWFSFAPPTWVILLALLAWMLQILFLGFNFRKNFGI 300  
Qy 296 GEGEERKQAAFOVITKQYRLLGPMSPAEKTVTVLVLLVLTFTREPFGPGMDTVFA 355  
Db 301 GEKQEQOQAAVCYVLTQEHRLGPMPTFAEKASILFVLVLLVLTFTREPFGFGNLAFF 360  
Qy 356 NEKGQSMASDGTVAIFISLVFIIPSKIPLGLMQDPKPKGLKAPPAIILTWVNDKMPWN 415  
Db 361 NAKGESWSDGTVAIFIGIINFILPSKFPGLTQDPENFGKLKAPGLGLDWTQKMPWN 420  
Qy 416 IVILGGGFALAKSEQGLSEWLDKLTPIQHTPPSATVILCLLIAIFTECTSNVATT 475  
Db 421 IVLLGGGYALAKSERSGLSEWLNKLTPLQSPVAPAIATILSLVATFTECTSNVATT 480  
Qy 476 TLFPIILASMAQAICLHPYVLMPCITLAASLAFMLPVATPNATVPNAIVFSPGGLKVSDBARAG 535  
Db 481 TIFPIILASMAQAICLHPYVLMPCITLATSALFMLPVATPNATVPNAIVFSGDLKVLDMARAG 540  
Qy 536 FLLNIIGVLTATLSINWSIPFKLDTFPPSAHNS--TSQCLINPNSNTPV 584  
Db 541 FLLNIIGVLTALAINSWGIPFLSLHSPSAQSNNTTAAQCLPSLANTTTP 590  
RESULT 3  
ABB82951  
ID ABB82951 standard; protein; 592 AA.

```

Db      241  APNLVLQGINSLFPQNGVNVNFASFPAFTMTWVILLALLLAWLQILFLGFNFRKNFGI 300
QY      296  GEGEERKQAAFOVKTQYRLLGPMSPAEKTVTVLVLLVVLVFTREPFGFGWDTVFA 355
Db      301  GEKMOEQQAAYCVIQTEHRLGPMTPAEKATISILFVILVLLVFTREPFGFLGWNLAFF 360
QY      356  NEKGOSMASDGTVAIFISLVMEIIPSKI PGLMODPKPKGKLKAPPAILLTWKTVDNKPWN 415
Db      361  NAKGESMWSVSDGTVAIFIGIIMFIIPSKFPGLTQDPENFGKLKAPLGLLDKWTNQKMPWN 420
QY      416  IVLLGGGFALAKGSESGLSGLSEWLGDKLTPLQHIPPSATAVILCLIAIAFTECTSNVATT 475
Db      421  IVLLGGGYALAKGSESGLSGLSEWLGDKLTPLQSVPAIAIILSLVATTECTSNVATT 480
QY      476  TFLPILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVSFGLKVSMDARAG 535
Db      481  TIFLPILASMAQAICLHPLYVMLPCTLATSLAFMLPVATPPNAIVSFGLKVLDMARAG 540
QY      536  FLNIIIGVLAITLSINSWSIPIFKLDTFSPWAHSN-TSQCLLNPSNSTVP 584
Db      541  FLNIIIGVLIILAINSWGIPFLSLHSFSPWAQSNNTAQCLPSLANTTTP 590

```

```

RESULT 4
ABR40100
ID  ABR40100 standard; protein; 592 AA.
XX
AC  ABR40100;
XX
DT  24-JUL-2003 (first entry)
XX
DE  Human renal sodium dicarboxylate co-transporter NADC1.
XX
KW  Anti-diabetic; anorectic; sodium dicarboxylate cotransporter; human;
KW  sodium tricarboxylate cotransporter; HepNadc; hepatocyte; HepG2;
KW  diabetes; obesity; lipid metabolism; aging; NADC1.
XX
OS  Homo sapiens.
XX
PN  WO2003029465-A1.
XX
PD  10-APR-2003.
XX
PF  27-SEP-2002; 2002WO-JP010038.
XX
PR  28-SEP-2001; 2001JP-00299433.
XX
PR  28-AUG-2002; 2002JP-00249016.
XX
PA  (SAKA ) OTSUKA PHARM CO LTD.
XX
PI  Kanemoto N, Omori Y, Sugano S, Obuchi Y;
XX
WI  WPI; 2003-354728/33.
XX
PT  Sodium di- or tricarboxylate cotransporter gene (HepNadc) for treatment
PT  and prevention of diabetes, obesity, for improving lipid metabolism and
PT  aging.
XX
PS  Example 1; Fig 1; 75pp; Japanese.
XX
CC  The present invention relates to human sodium di- or tricarboxylate
CC  cotransporter protein (HepNadc; ABR40097). HepNadc gene is expressed in
CC  human hepatocyte cancer-origin cells HepG2. Compounds which control the
CC  expression of the HepNadc gene and activity of the HepNadc polypeptide
CC  may be useful for treating and preventing diabetes, obesity, for
CC  improving lipid metabolism and aging. The present sequence is the NADC1
CC  protein (GenBank U26209), which was used in a sequence alignment with the
CC  HepNadc sequence
XX
SQ  Sequence 592 AA;

```

Query Match 78.2%; Score 2392; DB 6; Length 592;

```

Best Local Similarity 76.9%; Pred. No. 26-232;
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;

QY      1  MATCHPALWAYAYFYLLVILCLPIFLPLPLIVOTKAYCAYSIILMALLNCTEALPLAVTA 60
Db      1  MATCHQALWAYESYIIVFVFPILLPLPLVPSKEAYCAYAIILMALFWCTEALPLAVTA 60
QY      61  LPFIVLPFLMGIMDASEVCIEYFKDKNILFVGLGMVAIAVEHWNHKLRIALQVLLIIGVR 120
Db      61  LPFLILFPFMGIVDASEVAVEYVKDSNLLFFGGLLVIAIAVEHWNHKLRIALRVLLIIGVR 120
QY      121  PALLLGLFVLVTAFLSMWISNTATTAMVPIGHAVLEQLQSGK--KDVEGGNNNPFTELQ 178
Db      121  PAPLILGLFVLVTAFLSMWISNTATGAMVPIAHAVLDQLHSSQASSNVBEGSNNPFELQ 180
QY      179  EBCPOKEVTKLNGQ--PV-SAPSPRTOKTQEHHRFSQGLSICICYSASIGGIATLTGT 235
Db      181  EFSPOKEVTKLNGQALPVTSSASSSEGRAHLSQKHLHTQCMSLCVCYSASIGGIATLTGT 240
QY      236  TPNLVLQGVNSLFPQNGVNVNFASFPAFTMTIILLLALLLWLVQLVFLGVNFRKNFGF 295
Db      241  APNLVLQGINSLFPQNGVNVNFASFPAFTMTWVILLALLLWLVQLVFLGFNFRKNFGI 300
QY      296  GEGEERKQAAFOVKTQYRLLGPMSPAEKTVTVLVLLVVLVFTREPFGFGWDTVFA 355
Db      301  GEKMOEQQAAYCVIQTEHRLGPMTPAEKATISILFVILVLLVFTREPFGFLGWNLAFF 360
QY      356  NEKGOSMASDGTVAIFISLVMEIIPSKI PGLMODPKPKGKLKAPPAILLTWKTVDNKPWN 415
Db      361  NAKGESMWSVSDGTVAIFIGIIMFIIPSKFPGLTQDPENFGKLKAPLGLLDKWTNQKMPWN 420
QY      416  IVLLGGGFALAKGSESGLSGLSEWLGDKLTPLQHIPPSATAVILCLIAIAFTECTSNVATT 475
Db      421  IVLLGGGYALAKGSESGLSGLSEWLGDKLTPLQSVPAIAIILSLVATTECTSNVATT 480
QY      476  TFLPILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVSFGLKVSMDARAG 535
Db      481  TIFLPILASMAQAICLHPLYVMLPCTLATSLAFMLPVATPPNAIVSFGLKVLDMARAG 540
QY      536  FLNIIIGVLAITLSINSWSIPIFKLDTFSPWAHSN-TSQCLLNPSNSTVP 584
Db      541  FLNIIIGVLIILAINSWGIPFLSLHSFSPWAQSNNTAQCLPSLANTTTP 590

RESULT 5
ABR57025
ID  ABR57025 standard; protein; 592 AA.
XX
AC  ABR57025;
XX
DT  05-AUG-2003 (first entry)
XX
DE  Human NADC-1 amino acid sequence.
XX
KW  Human; TCHI69; dicarboxylate transport; hepatotrophic; cytostatic;
KW  nephrotropic; vasotropic; antidiabetic; liver disease; hepatitis;
KW  hepatic sclerosis; alcohol-related liver disease; prostate disease;
KW  prostatitis; prostatic hypertrophy; spleen disease; spleen hyperactivity;
KW  kidney disease; nephritis; kidney failure; nephritis; dropsey; diabetes;
KW  diabetes-associated renal disease; metabolic disease; hyperlipaemia;
KW  circulatory disease; arteriosclerosis; cancer; NADC-1.
XX
OS  Homo sapiens.
XX
PN  WO2003025168-A1.
XX
PD  27-MAR-2003.
XX
PF  13-SEP-2002; 2002WO-JP009444.
XX
PR  17-SEP-2001; 2001JP-00281992.
XX
PR  02-OCT-2001; 2001JP-00306873.
XX
PR  16-APR-2002; 2002JP-00113279.

```

XX (TAKA ) TAKEDA CHEM IND LTD.  
 XX Nakanishi A, Uno Y, Sagiya Y;  
 XX WPI; 2003-313352/30.  
 XX Protein TCH169 with dicarboxylate transport activity for treatment and  
 XX disorders.  
 XX Example 1; Fig 1-2; 132pp; Japanese.  
 XX The present invention describes protein TCH169 and its salts having  
 XX dicarboxylate transport activity. TCH169 has hepatotropic, cytostatic,  
 XX nephrotropic, vasotropic and antidiabetic activities. The TCH169 protein  
 XX and polynucleotide can be used in the treatment, prevention and diagnosis  
 XX of liver disease (such as hepatitis, hepatic sclerosis and alcohol-  
 XX related liver disease); prostate disease (such as prostatitis and  
 XX prostatic hypertrophy); spleen disease (such as spleen hyperactivity);  
 XX kidney disease (such as nephritis, kidney failure, nephritis, dropsy and  
 XX diabetes-associated renal disease); metabolic disease (such as diabetes);  
 XX circulatory disease (such as hyperlipaemia and arteriosclerosis); and  
 XX cancer (such as non-small cell lung cancer, liver cancer, renal cancer,  
 XX ovarian cancer, prostate cancer, stomach cancer, pancreatic cancer,  
 XX breast cancer, colon cancer, bladder cancer and womb cancer). The present  
 XX sequence represents a human NADC-1 amino acid sequence which is given in  
 XX comparison with the human TCH169 amino acid sequence in an example from  
 XX the present invention  
 XX Sequence 592 AA;  
 SQ

Query Match 78.2%; Score 2392; DB 6; Length 592;  
 Best Local Similarity 76.9%; Pred. No. 2e-232;  
 Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;

Qy 1 MATCPALWAYRYLVIVLCIFLPLPLIVOTKEAYCAYSIIMALLMCTEALPLAVTA 60  
 Db 1 MATCPALWAYRYLVIVLCIFLPLPLIVOTKEAYCAYSIIMALLMCTEALPLAVTA 60  
 Qy 61 LFPVLPFLMGIMDASEVICIEYFDNTNLFVGLMVAIVAEHNLKRIALQVLLIIVR 120  
 Db 61 LFPVLPFLMGIMDASEVICIEYFDNTNLFVGLMVAIVAEHNLKRIALQVLLIIVR 120  
 Qy 121 PALLLGLPMLVTAFLSMISNTATTAMVPRTGHVLEQLQSK--KQVEGGNNPTRELQ 178  
 Db 121 PALLLGLPMLVTAFLSMISNTATTAMVPRTGHVLEQLQSK--KQVEGGNNPTRELQ 180  
 Qy 179 EEPQKEVTKLDNGQ--PV-SAPSEPRTKQTEHRRFSQGLSLCICYASIGGIATLTGT 235  
 Db 181 EPSQKEVTKLDNGQALPVTSSASSEGRAHLQKHLHLTQCMSLCVCYASIGGIATLTGT 240  
 Qy 236 TPNLVLOQVNSLFPQNGVNVNFASFQFAPFTMIILLALLMQLVFLGVNFRKQNGF 295  
 Db 241 APNVLQSQINSLFPQNGVNVNFASFQFAPFTMIILLALLMQLVFLGVNFRKQNGF 300  
 Qy 296 GEGEERKQAFQVKTQYRLGPMSPAEKTVTLFVLLVLTFTREGRPGGCDTVFA 355  
 Db 301 GEGEERKQAFQVKTQYRLGPMSPAEKTVTLFVLLVLTFTREGRPGGCDTVFA 360  
 Qy 356 NEKQSMASDGTVAIFISLVNFIIPSKIPGLMQDPKPKGLKAPAILITKVTNDKMPWN 415  
 Db 361 NAKGESMVSDDTVAIFIGIINFIIPSKIPGLMQDPKPKGLKAPAILITKVTNDKMPWN 420  
 Qy 416 IVILLGGFALAKSGESGLSEWLDKLTPLQHPISPSATVILCLLIAIETECTSNVATT 475  
 Db 421 IVILLGGFALAKSGESGLSEWLDKLTPLQHPISPSATVILCLLIAIETECTSNVATT 480  
 Qy 476 TLFLPILASMAQAATCLHPLYVLMCTLAASLAFMLPVATPNVAIVFSGGLKVSMDARAG 535  
 Db 481 TLFLPILASMAQAATCLHPLYVLMCTLAASLAFMLPVATPNVAIVFSGGLKVSMDARAG 540  
 Qy 536 FLNIIIGVLAITLSINSWSIPIFKLDTFPFSAHNSN-TSQCLLNFSNSTVP 584

Db 541 FLNIIIGVLAITLSINSWSIPIFKLDTFPFSAHNSN-TSQCLLNFSNSTVP 590

RESULT 6  
 ADE63755  
 ID ADE63755 standard; protein; 592 AA.  
 XX ADE63755;  
 XX 29-JAN-2004 (first entry)  
 XX Human Protein Q13183, SEQ ID NO 9699.  
 XX Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX Homo sapiens.  
 OS WO2003016475-A2.  
 XX 27-FEB-2003.  
 XX 14-AUG-2002; 2002WO-US025765.  
 XX 14-AUG-2001; 2001US-0312147P.  
 XX 01-NOV-2001; 2001US-0346382P.  
 XX 26-NOV-2001; 2001US-0333347P.  
 XX (GEHO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 PA Woolf C, D'urso D, Befort K, Costigan M;  
 PI WPI; 2003-268312/26.  
 DR GENBANK; Q13183.  
 PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX Claim 1; Page; 1017pp; English.  
 CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 592 AA;  
 SQ

Query Match	78.2%;	Score 2392;	DB 7;	Length 592;	
Best Local Similarity	76.9%;	Pred. No. 2e-232;			
Matches	454;	Conservative	63;	Mismatches	67;
				Indels	6;
				Gaps	4;
QY	1	MATCWALWAYRYLIVLCPLPIFLPLPLIVQKEAYCAYSIIIMALLWCTEALPLAVTA	60		
DB	1	MATCWALWAYRYLIVLCPLPIFLPLPLIVQKEAYCAYSIIIMALLWCTEALPLAVTA	60		
QY	61	LFPIVLFLPLMGIMDASEVCIEYKDTNIFLVGGLMVAIAVEHWNHKLRIALQVLLIIGVR	120		
DB	61	LFPIVLFLPLMGIMDASEVCIEYKDTNIFLVGGLMVAIAVEHWNHKLRIALQVLLIIGVR	120		
QY	121	PALLLGFLMVTAFSLMWSINTATTAMVPIGHAVLEQLQGSK--KDVGGNNNPTFELQ	178		
DB	121	PAPLILGFLMVTAFSLMWSINTATTAMVPIGHAVLEQLQGSK--KDVGGNNNPTFELQ	180		
QY	179	EECPQKEVTKLDNGQ--PV-SAPSEPTOKTOEHHRFSQGLSLCICYSASIGGIATLTGT	235		
DB	181	EECPQKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHTQCMSCVCYSASIGGIATLTGT	240		
QY	236	TPNLVLQGVNSLFPQNGNVNFAFWGFAFPTMTIILLALLWQLVLFVGNFRKNFGF	295		
DB	241	APNLVLQGVNSLFPQNGNVNFAFWGFAFPTMTIILLALLWQLVLFVGNFRKNFGF	300		
QY	296	GEGERKQAAFOVIKTYRLLGPMGSAEKTIVTLFVLLVLTFTREPFGFGWGTVFA	355		
DB	301	GERKQAAFOVIKTYRLLGPMGSAEKTIVTLFVLLVLTFTREPFGFGWGTVFA	360		
QY	356	NEGQSWASDGTVAIFSLVNFIPSKI PGLMDDPKPKGKLKAPPAILTKTNDKMPWN	415		
DB	361	NAGGESMVSDGTVAIFSLVNFIPSKI PGLMDDPKPKGKLKAPPAILTKTNDKMPWN	420		
QY	416	IVLLGGGFALAKGSGSLSEWLGDKLTPLQHI PPSATAVILCLIAITFECTSNVATT	475		
DB	421	IVLLGGGFALAKGSGSLSEWLGDKLTPLQHI PPSATAVILCLIAITFECTSNVATT	480		
QY	476	TLFPLILASMAQAICLHPLVYMLPCTILASLAFMLPVATPPNAIVFSFGGLKTVSDMARAG	535		
DB	481	TIFLPLILASMAQAICLHPLVYMLPCTILASLAFMLPVATPPNAIVFSFGGLKTVSDMARAG	540		
QY	536	FLANIIGVLAITISINWSIPFKLDTFPPSWAHSN-TSQCLLNPSTVTP	584		
DB	541	FLANIIGVLAITISINWSIPFKLDTFPPSWAHSN-TSQCLLNPSTVTP	590		
RESULT 7					
ABM84184					
ID	ABM84184	standard; protein; 519 AA.			
XX	ABM84184;				
XX					
DT	18-NOV-2004	(first entry)			
XX					
DE		Human diagnostic and therapeutic pprotein SEQ ID NO:4433.			
XX		gene therapy; human diagnostic and therapeutic polynucleotide; dithp.			
KW					
XX					
OS		Homo sapiens.			
XX					
PN	WO2004023973-A2.				
XX					
PD	25-MAR-2004.				
XX					
PF	12-SEP-2003;	2003WO-US028227.			
XX					
PR	12-SEP-2002;	2002US-0410259P.			
PR	12-SEP-2002;	2002US-0410260P.			
XX					
PA	(INCY-) INCYTE CORP.				
XX					
PI	Schmidt JP, Wright RJ, Bruns CM, Marjanovic NM, Shen F;				
PI	Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;				
PI	Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;				

PI	Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;				
PI	Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;				
PI	Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;				
PI	Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;				
PI	Fatury S, Shi X, Suarez CU;				
XX					
DR	WPI; 2004-329368/30.				
DR	N-PSDB; ACN42836.				
XX					
PT	New diagnostic and therapeutic polynucleotides and polypeptides, useful				
PT	in diagnosing a condition, disease or disorder associated with human				
PT	molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or				
XX	in gene mapping.				
PS	Claim 27; Page; 190pp; English.				
XX					
CC	The invention relates to novel diagnostic and therapeutic polynucleotides				
CC	selected from one of the 2722 sequences defined in the specification. A				
CC	polynucleotide of the invention may have a use in gene therapy. The human				
CC	diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be				
CC	used to diagnose a particular condition, disease or disorder associated				
CC	with human molecules, e.g. cell proliferative disorders,				
CC	autoimmune/inflammatory disorder, developmental disorder, endocrine				
CC	disorder, neurological disorders, gastrointestinal disorders, or				
CC	infections caused by virus, bacteria, fungi or parasite. The dithp				
CC	molecules may also be used in genetic mapping, in identifying individuals				
CC	from minute biological samples, in detecting single nucleotide				
CC	polymorphisms, as molecular weight markers, and for somatic or germline				
CC	gene therapy. The present sequence represents a dithp protein of the				
CC	invention. Note: The sequence data for this patent is not represented in				
CC	the printed specification, but was obtained in electronic format directly				
CC	from WIPO at www.wipo.int/pct/en/sequences/listing.htm				
XX					
SQ	Sequence 519 AA;				
Query Match	67.8%;	Score 2073.5;	DB 8;	Length 519;	
Best Local Similarity	68.8%;	Pred. No. 2.9e-200;			
Matches	406;	Conservative	48;	Mismatches	57;
				Indels	79;
				Gaps	5;
QY	1	MATCHPALWAYRYLIVLCPLPIFLPLIVQKEAYCAYSIIIMALLWCTEALPLAVTA	60		
DB	1	MATCHPALWAYRYLIVLCPLPIFLPLIVQKEAYCAYSIIIMALLWCTEALPLAVTA	60		
QY	61	LFPIVLFLPLMGIMDASEVCIEYKDTNIFLVGGLMVAIAVEHWNHKLRIALQVLLIIGVR	120		
DB	61	LFPIVLFLPLMGIMDASEVCIEYKDTNIFLVGGLMVAIAVEHWNHKLRIALQVLLIIGVR	120		
QY	121	PALLLGFLMVTAFSLMWSINTATTAMVPIGHAVLEQLQGSK--KDVGGNNNPTFELQ	178		
DB	121	PAPLILGFLMVTAFSLMWSINTATTAMVPIGHAVLEQLQGSK--KDVGGNNNPTFELQ	180		
QY	179	EECPQKEVTKLDNGQ--PV-SAPSEPTOKTOEHHRFSQGLSLCICYSASIGGIATLTGT	235		
DB	181	EECPQKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHTQCMSCVCYSASIGGIATLTGT	240		
QY	236	TPNLVLQGVNSLFPQNGNVNFAFWGFAFPTMTIILLALLWQLVLFVGNFRKNFGF	295		
DB	241	APNLVLQGVNSLFPQNGNVNFAFWGFAFPTMTIILLALLWQLVLFVGNFRKNFGF	293		
QY	296	GEGERKQAAFOVIKTYRLLGPMGSAEKTIVTLFVLLVLTFTREPFGFGWGTVFA	355		
DB	294	-----	293		
QY	356	NEGQSWASDGTVAIFSLVNFIPSKI PGLMDDPKPKGKLKAPPAILTKTNDKMPWN	415		
DB	294	-----WVSDGTVAIFSLVNFIPSKI PGLMDDPKPKGKLKAPPAILTKTNDKMPWN	347		
QY	416	IVLLGGGFALAKGSGSLSEWLGDKLTPLQHI PPSATAVILCLIAITFECTSNVATT	475		
DB	348	IVLLGGGFALAKGSGSLSEWLGDKLTPLQHI PPSATAVILCLIAITFECTSNVATT	407		
QY	476	TLFPLILASMAQAICLHPLVYMLPCTILASLAFMLPVATPPNAIVFSFGGLKTVSDMARAG	535		

Db 408 TIFLPIILASMAQAICHLPLVYMLPCTLATSLAFMLPVATPPNAIVFSGDLKVLDMARAG 467

QY 536 FLNLIIGVLAITLSINSWSIPIFKLDTFPPSWAHNS-TSQCLLNPSNSTVP 584

Db 468 FLNLIIGVLAITLSINSWSIPIFKLDTFPPSWAHNS-TSQCLLNPSNSTVP 517

RESULT 8

ID ADH22543

XX ADH22543 standard; protein; 519 AA.

AC ADH22543;

XX 11-MAR-2004 (first entry)

DE Human transporter & ion channel (TRICH) protein SeqID41.

XX human; transporters and ion channel; TRICH; cell proliferative;

KW arteriosclerosis; cancer; autoimmune/inflammatory; AIDS; asthma;

KW neurological; epilepsy; stroke; developmental; Cushing's syndrome;

KW hypothyroidism; infection; gene therapy; cytostatic; antiinflammatory;

KW immunosuppressive; antiasthmatic; anticonvulsant; nootropic;

KW neuroprotective; single nucleotide polymorphism; SNP.

XX Homo sapiens.

OS WO2003093444-A2.

XX 13-NOV-2003.

XX 02-MAY-2003; 2003WO-US014026.

XX 03-MAY-2002; 2002US-0377435P.

PR 03-MAY-2002; 2002US-0377444P.

PR 05-JUN-2002; 2002US-0386497P.

PR 11-JUN-2002; 2002US-0388180P.

XX (INCY-) INCYTE CORP.

XX Baughn MR, Becha SD, Bulloch SA, Chang H, Elliott VS;

PI Embling BM, Griffin JA, Hafalata AJA, Ison CH, Jackson AA, Jiang X;

PI Jin P, Kable AE, Khare R, Lee SY, Lee S, Mason PM, Marquis JP;

PI Rankumar J, Richardson TW, Swarnakar A, Tran UK, Chawla NK;

PI Wilson AD;

XX WPI: 2004-022655/02.

DR N-PSDB; ADH22609.

XX New human transporters and ion channels (TRICH), useful for diagnosing,

PT treating and preventing diseases or conditions associated with the

PT aberrant TRICH expression e.g. cancer, AIDS, atherosclerosis, epilepsy,

PT or infections.

XX Claim 1; SEQ ID NO 41; 448pp; English.

XX This invention relates to novel isolated polynucleotides identified as

CC human transporters and ion channels (TRICH), and the encoded polypeptides

CC thereof. Specifically, it describes using these TRICH molecules, as well

CC as agonists, antagonists, antibodies, expression vectors and host cells,

CC in appropriate screening and toxicity assays to assess the effects of

CC exogenous compounds on TRICH expression. The present invention describes

CC TRICH compositions that are useful in the diagnosis, treatment and

CC prevention of various disorders such as cell proliferative (e.g.

CC arteriosclerosis, cancer), autoimmune/inflammatory (e.g. AIDS, asthma),

CC neurological (e.g. epilepsy, stroke) and developmental (e.g. Cushing's

CC syndrome hypothyroidism) and for infections. Accordingly, these TRICH

CC molecules can be used for gene therapy purposes and exhibit various

CC activities such as cytostatic, antiinflammatory, immunosuppressive,

CC antiasthmatic, anticonvulsant, nootropic and neuroprotective.

CC Furthermore, a microarray is useful in monitoring or measuring protein-

CC protein interactions, drug-target interactions and gene expression

CC profiles. This polypeptide sequence is a human TRICH protein of the

CC invention.

XX SQ Sequence 519 AA;

Query Match 67.5%; Score 2064.5; DB 8; Length 519;

Best Local Similarity 68.6%; Pred. No. 2.3e-199;

Matches 405; Conservative 48; Mismatches 58; Indels 79; Gaps 5;

QY 1 MATCPALWAYRFLYLIIVLCPLIFLLPLPLIVQTKYACAYSIILMALLWCTEALPLAVTA 60

Db 1 MATCQALWAYRSYLIIVFVILLPLPLIPVPSKEAYCAYAILMALLWCTEALPLAVTA 60

QY 61 LPPIVLPMLGIMDASEVICIEYFKDTNLFVGGMLVAIAVEHNNLHKRIALQVLIIGVR 120

Db 61 LPPLILFPMGIVDASEVAVEYKDSNLLFFGGLLVAIAVEHNNLHKRIALRVLLIVGR 120

QY 121 PALLILGFMVLTAFLSMWISNTATTAMVPIGHAIVLEOLQSGK--KDYEGGNNNPFELQ 178

Db 121 PAPILIGFMVLTAFLSMWISNTATYATSAWMPVIAHVDQLHSSQASSNVYEGSNNPFELQ 180

QY 179 EECQKEVTKLDNGQ--PV-SAPSEPRTKQTEHHRFSQGLSLCICYSASIGGIATLTGT 235

Db 181 EPSQKEVTKLDNGQALPVTSSASSEGRAHLQKHHLTQCHSLVCVCYSASIGGIATLTGT 240

QY 236 TPNLVLOGQVNSLPQNGNVNPFASWFGFAFPTMIILLLLAWMLQVLFGLVYNERKFGF 295

Db 241 APNLVLOGQVNSLPQNGNVNPFASWFGFAFPTMIILLLLAWMLQVLFGLVYNERKFGF 293

QY 296 GEGEERKQAAQVQIKTOYRLLGPMSPFAEKTIVTVLVLLVVLVFTREPFGPGWDTVFA 355

Db 294 ----- 293

QY 356 NEKGQSMASDGTVAIFISLVNFIIPSKI PGLMQDPKPKGLKAPPAITLTKVNDKMPWN 415

Db 294 -----WVSDGTVAIFIGIMFIIPSKFPGTQDPENPGKLPGLGLDWDKIVNQKMPWN 347

QY 416 IVILLGGFALAKGSEQSLSEWLGDKLTPLQHIPPSATAVILCLLIAIFTCTSNVATT 475

Db 348 IVLLGGYALAKGSESGSLSEWLGDKLTPLQSVPAIPAIAIISLVVATFECTSNVATT 407

QY 476 TLFPLILASMAQAICHLPLVYMLPCTLATSLAFMLPVATPPNAIVFSGDLKVLDMARAG 535

Db 408 TIFLPIILASMAQAICHLPLVYMLPCTLATSLAFMLPVATPPNAIVFSGDLKVLDMARAG 467

QY 536 FLNLIIGVLAITLSINSWSIPIFKLDTFPPSWAHNS-TSQCLLNPSNSTVP 584

Db 468 FLNLIIGVLAITLSINSWSIPIFKLDTFPPSWAHNS-TSQCLLNPSNSTVP 517

RESULT 9

ADG16995

ID ADG16995 standard; protein; 619 AA.

XX ADG16995;

XX 26-FEB-2004 (first entry)

DE African clawed frog transporter protein.

XX african clawed frog; transporter protein; pharmacogenomic analysis.

XX Xenopus laevis.

XX US2003186381-A1.

XX 02-OCT-2003.

XX 12-MAY-2003; 2003US-00435631.

XX 13-JUN-2000; 2000US-0211220P.

XX 05-DEC-2000; 2000US-00729094.

XX (APPL-) APPLERA CORP.

PI Chaturvedi K, Wei M, Ketchum KA, Difrancesco V, Beasley EM;  
XX WPI; 2004-032009/03.  
XX New isolated peptide for identifying its modulator, or an agent that  
PT binds to it, where the agent is used in a pharmaceutical composition for  
PT treatment of a disease or condition mediated by a human transporter  
PT protein.  
XX Disclosure; SEQ ID NO 4; 68pp; English.  
XX The invention relates to an isolated transporter protein. The invention  
CC is used for identifying its modulator, or an agent that binds to it,  
CC where agent is useful for a pharmaceutical composition for treatment of a  
CC disease or condition mediated by a human transporter protein. It is used  
CC as models for the development of human therapeutic targets; aid in the  
CC identification of therapeutic proteins; and serves as targets for the  
CC development of human therapeutic agents that modulate transporter  
CC activity, in cells and tissues that express the transporter. It is useful  
CC for development of commercially important products and services; and in  
CC pharmacogenomic analysis. The invention allows effective clinical design  
CC of treatment compounds and dosage regimens. The present sequence  
CC represents the amino acid sequence of an african clawed frog transporter  
CC protein.  
XX Sequence 619 AA;  
SQ  
Query Match 64.5%; Score 1971.5; DB 8; Length 619;  
Best Local Similarity 63.2%; Pred. No. 7.9e-190;  
Matches 386; Conservative 87; Mismatches 85; Indels 53; Gaps 10;  
QY 10 AYRFYLIVLCPLPIFLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTALFPVILFPL 69  
DB 10 ANRNYFIIFLPLPLPLVPLVFTKASCGFVIIVMALFWCTEALPLAVTALFPVILFPM 69  
QY 70 MGIMDASEVCIEYFKDNLIFVGLMVAIAVEHNLHKLRIALQVLLIIGVRPALLLLGFM 129  
DB 70 MGIMDSTAVCSQYKLDNMLFVGLMVAIAVEHNLHKLRIALQVLLIIGVRPALLLLGFM 129  
QY 130 LVTAFLSMWISNTATTAMVPVIGHAVLEQLQSGK----KQVEGNN----- 171  
DB 130 VVTAFLSMWISNTATTAMVPVIAQVMEQLHSEGVKDERVGNSTQKNVGNWMDMYE 189  
QY 172 -----NPTFEOBEC-----POKEVTKLDNGQPVSAPEPRTQKT 206  
DB 190 SVMPSGKMAIAIDYATENEGEIEQKSTKDEPSKQ-EKQSIGIVLIEPEDEKTEEK 248  
QY 207 Q--EHRFSGQLSLCYCSASIGGIATLTGTTPLNLVQGVNSLFPQNGVNVNFAWFGF 264  
DB 249 QKEKHLKICKGMSLVCYCSASIGGIATLTGTTPLNLVQGVNSLFPQNGVNVNFAWFGF 308  
QY 265 APTNMIILLLLAWLQVLEFLGVNFRKNFGFGEGEE--ERKQAAFOVIKTOYELLGPMSP 322  
DB 309 APTNMIILLLLAWLQVLEFLGVNFRKNFGFGEGEE--ERKQAAFOVIKTOYELLGPMSP 368  
QY 323 AEKTVTVLVLLVLFVFTREPGFPGMGDTVFANEGQSMASDGTVAIFISLVNMFIPSK 382  
DB 369 AEISLVLFLLVLLVFTREPGFPGMGWA-TISFNKGGKEMWTATVAIFVSLNMFEPSE 427  
QY 383 IPGL-MQDPKKPG---KLKAPPAITLTKTVNDKMPNIVILLGGGFALAKGSGSLSEW 438  
DB 428 LPSFKYQDTPKGMKPKRLRVPVALLDWTNKNMPNIVILLGGGFALAKGSGSLSLW 487  
QY 439 LGKDLPLQHIPSAATAVILICLIAIFTECTSNVATTTPLPLASNAQAICLHPVYML 498  
DB 488 LGKDLPLQHIPSAATAVILICLIAIFTECTSNVATTTPLPLASNAQAICLHPVYML 547  
QY 499 PCTLAASLAFMLPVATPPNIAVFSFGGLKYSVMARAGFLNIIGVLAITLSINSWSPIF 558  
DB 548 PCTLAASLAFMLPVATPPNIAVFSFGGLKYSVMARAGFLNIIGVLAITLSINSWSPIF 607  
QY 559 KLDTPFPSSWAHS 569  
| | | | | | | | | |

DB 608 NLGTFFPSWANA 618  
RESULT 10  
ADP64799  
ID ADP64799 standard; protein; 581 AA.  
XX AC ADP64799;  
XX DT 26-AUG-2004 (first entry)  
XX Zebrafish Na+-coupled citrate transporter protein.  
DE sodium-coupled citrate transporter; transmembrane citrate transporter;  
KW life span; weight reduction; weight gain prevention; blood cholesterol;  
KW triglyceride; low density lipoprotein; glucose; obesity;  
KW hyperlipidemia; hypercholesterolemia; INDI protein.  
XX Danio rerio.  
OS WO2004048925-A2.  
PN 10-JUN-2004.  
PD 20-NOV-2003; 2003WO-US037054.  
PF 22-NOV-2002; 2002US-0428469P.  
PR 01-APR-2003; 2003US-0459441P.  
XX (MEDI-) MEDICAL COLLEGE GEORGIA RES INST.  
PA (GANA/) GANAPATHY V.  
PA (INOU/) INOUE K.  
PA (FEIY/) FEI Y.  
PI Ganapathy V, Inoue K, Fei Y;  
XX WPI; 2004-460797/43.  
DR N-PSDB; ADP64798.  
XX New isolated polynucleotide encoding a Na+-coupled citrate transporter  
PT (NaCT) polypeptide, useful as a drug target for the treatment of obesity,  
PT hyperlipidemia, and hypercholesterolemia.  
XX Claim 42; SEQ ID NO 12; 186pp; English.  
XX The invention relates to novel Na+-coupled citrate transporter proteins  
and their encoding genes. Inhibitors of transmembrane citrate  
transporters are useful for extending the lifespan, reducing weight,  
preventing weight gain or lowering blood cholesterol, triglyceride, LDL  
or glucose levels in a subject. The NaCT polypeptide is useful as a drug  
target for the treatment of obesity, hyperlipidemia, and  
hypercholesterolemia. This sequence corresponds to the zebrafish Na+-  
coupled citrate transporter protein.  
XX Sequence 581 AA;  
SQ  
Query Match 59.8%; Score 1829; DB 8; Length 581;  
Best Local Similarity 60.4%; Pred. No. 1.9e-175;  
Matches 349; Conservative 92; Mismatches 113; Indels 24; Gaps 6;  
QY 8 LWAYRFLVLCPLPIFLPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTALFPVILF 67  
DB 9 VVKMKNLTILFCTPFLPLPLVIGSKEAGCAVYVVLMAVYVCTEVLPLVATALLPAVL 68  
QY 68 PLMGINDASEVCIEYFKDNLIFVGLMVAIAVEHNLHKLRIALQVLLIIGVRPALLLLG 127  
DB 69 PLRIMESQDVCQYKLDNMLFVGLMVAIAVEHNLHKLRIALQVLLIIGVRPALLLLG 128  
QY 128 FMLVTAFLSMWISNTATTAMVPVIGHAVLEQLQSGKQVEGG---NNNPTFLEQECPOK 184  
DB 129 FMCVTAFLSMWISNTATTAMVPVIGHAVLEQLQSGKQVEGG---NNNPTFLEQECPOK 188  
QY 185 EVTKLDNGQPVSAPEP--RTQKTQEHHRFSGQLSLCYCSASIGGIATLTGTTPLNLVQ 242

Db 189 EKVL-NGDNFWSDEPESHSREERLWMSKGLTLCVYASIGTATLTGTGPNVLUM 247  
Qy 243 GQNSLFPQNGVNVNPFASFGAPFPMIILLALLMLQLVFLGVNFRKNFGFGEGER 302  
Db 248 GQMSQLFPNDPDIINPFASFGAPFPMIIMLTALMLQLVFLGVNFKKGTGKTEK 307  
Qy 303 KQAAFOVITQVRLGPNMFAEKTIVLVLVLLWFTREPGFFPGWGTVPANEKGQSM 362  
Db 308 EIAANVKEERHSLGPMFTGSLVLLFLLVLMFTDRDPGVDCWATRRFNADK--EF 365  
Qy 363 ASDGTVAIFISLWMTIPISKIPGLM-----QDPKKPKGLKAPPAIITKTVNDK 411  
Db 366 VDATVAVFVALLFVFPKPRCLCWRTESDTPVQSGSP-----TPALLTKVTKQK 420  
Qy 412 MPWNIVILGGGFALAKSGEQSLSEWLGDKLTPLQHIPPSATAVILCILIAIFTECTSN 471  
Db 421 MPWSIILLGGGFALAKSGEISGLSKWLGQSPQSQIPPMIAIVICLMIAITFECTSN 480  
Qy 472 VATTILFLPILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSFGLKVS DM 531  
Db 481 VATATLFLPILASMSQISQVNPPLYVMPCTLSASFAPMLPVATPPNAIVFSYGLKVS DM 540  
Qy 532 ARAGELLATIGVLAITLSINSHSIPFKLDTFSPWAHS 569  
Db 541 AKTGIVNNIIGLITSLTAINSGRAIFSLDTFPSWANT 578  
  
RESULT 11  
ABU65064  
ID ABU65064 standard; protein; 616 AA.  
XX AC ABU65064;  
XX DT 20-MAY-2003 (first entry)  
XX DE Human NOVI4c protein.  
XX KW NOVI; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;  
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;  
KW human.  
XX OS Homo sapiens.  
XX PN WO200272757-A2.  
XX PD 19-SEP-2002.  
XX PF 08-MAR-2002; 2002WO-US006908.  
XX PR 08-MAR-2001; 2001US-0274101P.  
PR 08-MAR-2001; 2001US-0274194P.  
PR 08-MAR-2001; 2001US-0274281P.  
PR 09-MAR-2001; 2001US-0274322P.  
PR 09-MAR-2001; 2001US-0274849P.  
PR 12-MAR-2001; 2001US-0275235P.  
PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276776P.  
PR 19-MAR-2001; 2001US-0276994P.  
PR 20-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 21-MAR-2001; 2001US-0277791P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278999P.  
PR 27-MAR-2001; 2001US-0279036P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0277338P.  
  
PR 30-MAR-2001; 2001US-0279995P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 04-APR-2001; 2001US-0280900P.  
PR 04-APR-2001; 2001US-0281194P.  
PR 13-APR-2001; 2001US-0283675P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 03-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
PR 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
PR 18-OCT-2001; 2001US-0330380P.  
PR 31-OCT-2001; 2001US-0335301P.  
PR 14-NOV-2001; 2001US-0332172P.  
PR 14-NOV-2001; 2001US-0332271P.  
PR 14-NOV-2001; 2001US-033272P.  
PR 14-NOV-2001; 2001US-0333184P.  
PR 21-NOV-2001; 2001US-0333272P.  
PR 03-DEC-2001; 2001US-0332094P.  
PR 03-DEC-2001; 2001US-0337426P.  
PR 04-DEC-2001; 2001US-0338092P.  
PR 03-JAN-2002; 2001US-0337185P.  
PR 07-MAR-2002; 2002US-0345705P.  
PR 07-MAR-2002; 2002US-00092900.  
XX (CURA-) CURAGEN CORP.  
XX Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;  
XX Zehrhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;  
XX Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;  
XX Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;  
XX Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;  
XX Lepley DM, Rieger DK;  
XX MPI; 2002-723332/78.  
XX N-FSDB; ABX97031.  
XX NOVI polypeptides and polynucleotides, useful for preventing or treating  
XX a disorder associated with aberrant NOVI expression or activity e.g.,  
XX cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
XX asthma.  
XX Claim 1; Page 134; 1103pp; English.  
XX This invention describes novel human NOVI polypeptides which have  
XX cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive  
XX activity. Pharmaceutical compositions comprising the NOVI proteins or  
XX nucleic acid molecules or NOVI antibodies are useful for preventing or  
XX treating a disorder associated with aberrant NOVI expression or activity  
XX e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
XX asthma. The products of the invention can be used for gene therapy or in  
XX a vaccine. ABU65041-ABU65218 represent the NOVI polypeptides encoded by  
XX ABX97008-ABX97185  
XX Sequence 616 AA;  
XX Query Match 53.2%; Score 1628; DB 5; Length 616;

Best Local Similarity 51.3%; Pred. No. 4.6e-155;		
Matches 314; Conservative 113; Mismatches 139; Indels 46; Gaps 7;		
Qy	1	MATCWPAALWAYRYLVLCPIPIELPLPLIVQKQKAYCAYSIIMALLWCTEALPLAVTA 60
Db	1	MASALSYVSKFSGFVLTFTPLLLPLVILMPAKVSCCAYVILMAIYWCTEVPLAVTS 60
Qy	61	LFPVLVPLMGIMDASVCIYEKFDKNILFVGGIMVAIAVEHNLKRIALQVLLIIGVR 120
Db	61	LMPVLVPLFQIILDSROVCVQYKQDNMLFGLGIIIVAVAVERNLKHRIALRILLWVGAK 120
Qy	121	PALLLGFMILVTAFLSMWISNTATTAMVPIGHAVLEQLQSGKQVEGG-----NNN 172
Db	121	PARMLGFMGVGTALLSMWISNTATTAMVPIVEAILQMEATSAATEAGLEGQGTINNLL 180
Qy	173	PTPE-----LQECQDQ-----KEVTKL-----DNGQVPSAPSEP 201
Db	181	NALEDDTVKAVLGKCKVAIISTYVKVEKLQINNMTPLKKLEKQEQDGLGPIRQDSA 240
Qy	202	RTKQTOEHHRFSOGLSLICIVSASIGGIATLTGTPNLVLOGOVNSLFPONGNVNFAFW 261
Db	241	QCQEDQBRKRLCKAMTILCICVASIGGTALTGTGPNVLLGQWNLFPDSKDLVNFASW 300
Qy	262	FGFAFTMIILLIALLWLQVLGLNFRKNGFGEGEERKQAAFOVIKTYRLILGPMWS 321
Db	301	FATFAPNMLVMLLFAWLQFVVMFSFKSGWGGLESKKNEKAALKVLQEEYVKLGPLS 360
Qy	322	FAKTVTVLVLLVLTFTREPGFFQWGTVPANEKGQSMASDGTVAIFISLVMFIIPS 381
Db	361	FAEINVLLICFLLVILFWSRDPGFMGPGLTVAWVEGETKSV-SDATVAIFVATLLFIIPS 419
Qy	382	KIP-----GLMDPKPKGKLKAPAILTWTVDKMPNIVILLGGFALAKSEOSGLS 436
Db	420	QKAPNFRSQTEGKSP-VLIAPPPLDKNVQEKVPGVILVLLGGFALAKSEASGLS 478
Qy	437	EWLGDKLTPHQHPPSATVILCLLIAIFTECTSNVATTLFLPILASMAQAICLHPLYV 496
Db	479	VWMSQMEPLHVAVPPAAITLILSLVAVTECTSNVATTLFLPILASMSRSIGLNPLYI 538
Qy	497	MLPCTLAASAFMLPVATPPNAIVFSGGLKVSMDARAGFLNIIIGVLAITLINSWSIP 556
Db	539	MLPCTLSASFAMLPVATPPNAIVFTYGHKLVADMVKTGVMNIIIGVFCVFLAVNTWGRA 598
Qy	557	IFKLDTPPSNAH 568
Db	599	IFDLDFPDWAN 610
RESULT 12		
ABU65063		
ID	ABU65063 standard; protein; 616 AA.	
XX	ABU65063;	
AC	(first entry)	
DT	20-MAY-2003	
XX	Human NOV14b protein.	
DE	NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;	
KW	hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;	
KW	human.	
XX	Homo sapiens.	
OS	WO200272757-A2.	
FN	19-SEP-2002.	
XX	08-MAR-2002; 2002WO-US006908.	
XX	08-MAR-2001; 2001US-0274101P.	
PR	08-MAR-2001; 2001US-0274194P.	
PR	08-MAR-2001; 2001US-0274281P.	

PR	08-MAR-2001; 2001US-0274322P.	
PR	09-MAR-2001; 2001US-0274849P.	
PR	12-MAR-2001; 2001US-0275235P.	
PR	13-MAR-2001; 2001US-0275578P.	
PR	13-MAR-2001; 2001US-0275579P.	
PR	13-MAR-2001; 2001US-0275601P.	
PR	14-MAR-2001; 2001US-0276000P.	
PR	16-MAR-2001; 2001US-0276776P.	
PR	19-MAR-2001; 2001US-0276994P.	
PR	20-MAR-2001; 2001US-0277239P.	
PR	20-MAR-2001; 2001US-0277321P.	
PR	20-MAR-2001; 2001US-0277327P.	
PR	21-MAR-2001; 2001US-0277791P.	
PR	22-MAR-2001; 2001US-0277833P.	
PR	23-MAR-2001; 2001US-0278152P.	
PR	26-MAR-2001; 2001US-0278894P.	
PR	27-MAR-2001; 2001US-0278999P.	
PR	27-MAR-2001; 2001US-0279036P.	
PR	28-MAR-2001; 2001US-0279344P.	
PR	30-MAR-2001; 2001US-0277338P.	
PR	30-MAR-2001; 2001US-0279959P.	
PR	30-MAR-2001; 2001US-0280233P.	
PR	02-APR-2001; 2001US-0280802P.	
PR	02-APR-2001; 2001US-0280822P.	
PR	02-APR-2001; 2001US-0280900P.	
PR	04-APR-2001; 2001US-0281194P.	
PR	13-APR-2001; 2001US-0283675P.	
PR	30-APR-2001; 2001US-0287424P.	
PR	02-MAY-2001; 2001US-0288066P.	
PR	03-MAY-2001; 2001US-0288342P.	
PR	03-MAY-2001; 2001US-0288528P.	
PR	15-MAY-2001; 2001US-0291190P.	
PR	16-MAY-2001; 2001US-0291099P.	
PR	16-MAY-2001; 2001US-0291240P.	
PR	30-MAY-2001; 2001US-0294485P.	
PR	31-MAY-2001; 2001US-0294889P.	
PR	31-MAY-2001; 2001US-0294899P.	
PR	18-JUN-2001; 2001US-0299027P.	
PR	19-JUN-2001; 2001US-0299303P.	
PR	19-JUN-2001; 2001US-0299310P.	
PR	10-JUL-2001; 2001US-0304354P.	
PR	31-JUL-2001; 2001US-0309198P.	
PR	16-AUG-2001; 2001US-0312903P.	
PR	10-SEP-2001; 2001US-0318462P.	
PR	12-SEP-2001; 2001US-0318770P.	
PR	27-SEP-2001; 2001US-0325430P.	
PR	27-SEP-2001; 2001US-0325681P.	
PR	18-OCT-2001; 2001US-0330380P.	
PR	31-OCT-2001; 2001US-0335301P.	
PR	14-NOV-2001; 2001US-0332172P.	
PR	14-NOV-2001; 2001US-0332712P.	
PR	14-NOV-2001; 2001US-0332722P.	
PR	14-NOV-2001; 2001US-0333184P.	
PR	14-NOV-2001; 2001US-0333272P.	
PR	21-NOV-2001; 2001US-0332094P.	
PR	03-DEC-2001; 2001US-0337426P.	
PR	03-DEC-2001; 2001US-0338092P.	
PR	04-DEC-2001; 2001US-0337185P.	
PR	03-JAN-2002; 2002US-0345705P.	
PR	07-MAR-2002; 2002US-00092900.	
XX	(CURA-) CURAGEN CORP.	
PA	Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;	
XX	Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;	
PI	Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;	
PI	Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;	
PI	Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;	
PI	Lepley DM, Rieger DK;	
XX	WFI; 2002-723332/78.	
DR	N-PSDB; ABX97030.	
XX		

PT NOVX polypeptides and polynucleotides, useful for preventing or treating  
PT a disorder associated with aberrant NOVX expression or activity e.g.,  
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
PT asthma.

PS Claim 1; Page 134; 1103pp; English.

XX This invention describes novel human NOVX polypeptides which have  
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive  
CC activity. Pharmaceutical compositions comprising the NOVX proteins or  
CC nucleic acid molecules or NOVX antibodies are useful for preventing or  
CC treating a disorder associated with aberrant NOVX expression or activity  
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
CC asthma. The products of the invention can be used for gene therapy or in  
CC a vaccine. AB065041-AB065218 represent the NOVX polypeptides encoded by  
CC ABX97008-ABX97185

XX Sequence 616 AA;

Query Match 53.2%; Score 1628; DB 5; Length 616;  
Best Local Similarity 51.3%; Pred. No. 4.6e-155;  
Matches 314; Conservative 113; Mismatches 139; Indels 46; Gaps 7;

QY 1 MATCHPALWAYRFLVILCLPIFLPLPLIVQTKAYCAYSIIIMALLWCTEALPLAVTA 60  
DB 1 MASALSYSVKFSFVILFTVTLPLLPVILMPAKVSCCAVYIIMAIYWCVEIPLAVTS 60

QY 61 LPIVLPFLMIGMDASEVCIEYFKDTNIIIVGGLMVAIAVEHNLHKKRIALQVLLIIGVR 120

DB 61 LMPVLLPFLFQILDRCQVCQYKMDTNMLFLGLLIVAVAVERNHKKRIALTLLWVCAK 120

QY 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLQGSKKDVEGG-----NNN 172

DB 121 PARMLGFMGVTAFLSMWISNTATTAMVPIVEAILQMEATSAATEAGLEGQGTINN 180

QY 173 PTFE-----LQECPO-----KEVTKL-----DNGQPVSAPEP 201

DB 181 NAEEDDTTKAVLGKCKVAILSTYVYKVEKLIQINNMTPLKKLEKOEQDGLGFGIRPQSA 240

QY 202 RTQKQEHRRFSQGLSLCICYASISGGIATLTGTPNLVLOGVNSLPPONGNVNFWASW 261

DB 241 QCQEQERKRLCKAMTLCICYAASISGGTATLTGTGNVVLQGMELPDSKDLNFWASW 300

QY 262 GFAPFTMIILLLAWMLQVLFLGNFRKNFGEGEBEERKQAAQVQIKTYRLLGPM 321

DB 301 FAFAPFNMLVMLFAWMLQVYMFSSFKSGGCGLESKNEKAAALKVLEBYRKLPLS 360

QY 322 FAEKTVTVLVLLVLTREPGFPFGDTPANEKGSNASDGTVAIFISLWFIIPS 381

DB 361 FAEINVLICFFLLVILWFSRDRGFMGLTVAWVEGETKSV-SDATVAIFVATLLFVPS 419

QY 382 KIP-----GLMQDPKPKGKLKAPPAILTKVTNDKMPNIVILGGFPALAKSGSOSGLS 436

DB 420 QPKPKNFRSQTEEGKSP-VLIAPPDLQWKVQEKVPMGIVLLGGFPALAKSGSASGLS 478

QY 437 EWLGDKLTPLOHIPPASATVILCLLIAIFTECTSNVATTTTLFPLTASMAQICLHPLYV 496

DB 479 VMGQMEPLHAPVPAATLILSLVAVTECTSNVATTTTLFPLFASMSRISGLNPLYI 538

QY 497 MLPCTLASLAPMLPVATPPNVAIVPSFGGLKVSMDARAGFLNIIGVLAITLINSWIP 556

DB 539 MLPCTLSASFAPMLPVATPPNVAIVFTYGHKLVADMVKTGVIMNIIGVFCVFLAVNTWGA 598

QY 557 IPKLDTTPSWAH 568

DB 599 IFDLDDHFDWAN 610

RESULT 13

ADK51048 standard; protein; 616 AA.

XX ADK51048;

XX 17-JUN-2004 (first entry)  
XX Human NOV18A protein sequence SeqID68.  
XX cytostatic; NOVX-agonist; NOVX-antagonist; vaccine; gene therapy; cancer;  
XX chromosome mapping; human; NOV18A.  
XX Homo sapiens.  
XX WO2003083046-A2.  
XX 09-OCT-2003.  
XX 01-APR-2003; 2003WO-US010142.  
XX 02-APR-2002; 2002US-00115479.  
XX 05-APR-2002; 2002US-0370349P.  
XX 08-APR-2002; 2002US-0370969P.  
XX 12-APR-2002; 2002US-0372019P.  
XX 22-APR-2002; 2002US-0374379P.  
XX 30-MAY-2002; 2002US-0384543P.  
XX 03-JUN-2002; 2002US-00160619.  
XX 15-AUG-2002; 2002US-0403748P.  
XX 04-NOV-2002; 2002US-00287226.  
XX 31-MAR-2003; 2003US-00403161.  
XX (CURA-) CURAGEN CORP.

XX Anderson DW, Bento P, Boldog FL, Burgess CE, Casman SJ, Furtak K;  
XX Gorman L, Gould-Rothberg BE, Gunther E, Heyes MP, Li L, Spytek KA;  
XX Stone DJ, Zhong M, Malyankar UM, Edinger SR, Patturajan M;  
XX Rothenberg ME, Smithson G;  
XX WPI; 2003-812539/76.  
XX N-PSDB; ADK51047.

XX New NOVX polypeptide, useful for preparing a composition for treating or  
XX preventing e.g. cancer or for chromosome mapping.  
XX Claim 1; SEQ ID NO 68; 433pp; English.

XX This invention relates to novel isolated polypeptides and the DNA  
XX sequences which encode them. The invention may be useful for the  
XX development of compounds with a cytostatic activity (as NOVX-agonists or  
XX antagonists) or vaccines. In addition, the disclosed sequences may be  
XX useful for gene therapy. The polypeptide is useful for preparing a  
XX composition for treating or preventing a pathological state in a mammal,  
XX for example cancer or for chromosome mapping. The present sequence is  
XX that of a human NOVX protein of the invention.

XX Query Match 53.2%; Score 1628; DB 7; Length 616;  
XX Best Local Similarity 51.3%; Pred. No. 4.6e-155;  
XX Matches 314; Conservative 113; Mismatches 139; Indels 46; Gaps 7;

QY 1 MATCHPALWAYRFLVILCLPIFLPLPLIVQTKAYCAYSIIIMALLWCTEALPLAVTA 60  
DB 1 MASALSYSVKFSFVILFTVTLPLLPVILMPAKVSCCAVYIIMAIYWCVEIPLAVTS 60

QY 61 LPIVLPFLMIGMDASEVCIEYFKDTNIIIVGGLMVAIAVEHNLHKKRIALQVLLIIGVR 120

DB 61 LMPVLLPFLFQILDRCQVCQYKMDTNMLFLGLLIVAVAVERNHKKRIALTLLWVCAK 120

QY 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLQGSKKDVEGG-----NNN 172

DB 121 PARMLGFMGVTAFLSMWISNTATTAMVPIVEAILQMEATSAATEAGLEGQGTINN 180

QY 173 PTFE-----LQECPO-----KEVTKL-----DNGQPVSAPEP 201

DB 181 NAEEDDTTKAVLGKCKVAILSTYVYKVEKLIQINNMTPLKKLEKOEQDGLGFGIRPQSA 240





PA (MILL/) LI L.  
PA (ZERH/) ZERHUSEN B D.  
PA (GUSE/) GUSEV V Y.  
PA (JIW/) JI W.  
PA (GORM/) GORMAN L.  
PA (MILL/) MILLER C E.  
PA (KEKU/) KEKUDA R.  
PA (PATT/) PATTURAJAN M.  
PA (GANG/) GANGOLLI E A.  
PA (VERN/) VERNET C A M.  
PA (GUOX/) GUO X S.  
PA (TCHE/) TCHERNEV V T.  
PA (FERN/) FERNANDES E R.  
PA (CASM/) CASMAN S J.  
PA (MALY/) MALYANKAR U M.  
PA (GERL/) GERLACH V.  
PA (LIUY/) LIU Y.  
PA (ANDE/) ANDERSON D W.  
PA (SPAD/) SPADERNA S K.  
PA (CATT/) CATTERTON E.  
PA (LEIT/) LEITE M W.  
PA (ZHON/) ZHONG H.  
PA (ALSO/) ALSOBROOK J P.  
PA (LEPL/) LEPLEY D M.  
PA (RIEG/) RIEGER D K.  
PA (BURG/) BURGESS C E.  
XX  
PI Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;  
PI Zerrhusen BD, Gusev VY, Ji W, Gorman L, Miller CE, Kekuda R;  
PI Pattrajan M, Gangolli EA, Vernet CAM, Guo XS, Tchernev VT;  
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;  
PI Anderson DW, Spaderna SK, Catterton E, Leite MW, Zhong H;  
PI Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;  
XX  
DR WPI; 2004-225693/21.  
DR N-PSDB; ADN61776.  
XX  
PT New NOVX polypeptides and nucleic acid molecules useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. cancer, diabetes,  
PT infection or obesity, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
PS Claim 1; SEQ ID NO 46; 786pp; English.  
XX  
CC The invention relates to an isolated polypeptide (designated NOVX, or  
CC NOV1-NOV127) comprising a sequence selected from 178 fully defined amino  
CC acid sequences (and their mature forms, variants and fragments). Also  
CC included are an isolated nucleic acid molecule encoding NOVX, a vector  
CC comprising the nucleic acid, a cell comprising the vector, methods for  
CC determining the presence or amount of the polypeptide or the nucleic acid  
CC molecule in a sample, methods for determining the presence of or  
CC predisposition to a disease associated with altered levels of expression  
CC of the above polypeptide or nucleic acid molecule in a first mammalian  
CC subject, a method for identifying an agent that binds to the above  
CC polypeptide, a method for identifying a potential therapeutic agent for  
CC use in the treatment of a pathology that is related to aberrant  
CC expression or physiological interactions of the polypeptide, a method of  
CC screening for a modulator of activity or of latency or predisposition to  
CC a pathology associated with the polypeptide and a method for modulating  
CC the activity of the polypeptide cited above. The composition and methods  
CC are useful for diagnosing, preventing or treating diseases such as  
CC diabetes, obesity, infectious diseases, anorexia, cancer-associated  
CC cachexia, cancer, neurodegenerative disorders like Alzheimer's disease or  
CC Parkinson's disease, immune disorders, haematopoietic disorders,  
CC dyslipidaemias, and other chronic diseases. These may also be used in  
CC chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The polypeptides are also useful as vaccines. The  
CC present sequence represents a NOVX protein of the invention.  
XX  
SQ Sequence 616 AA;

Query Match 53.2%; Score 1628; DB 8; Length 616;  
Best Local Similarity 51.3%; Pred. No. 4.6e-155;

Matches 314; Conservative 113; Mismatches 139; Indels 46; Gaps 7;  
QY 1 MATCPALWAFYFLIVLCLPIFLPLPLIVQTKEAYCAYSTILMALLWCTEALPATA 60  
DB ||: : : : : ||: : : : : ||: : : : : ||: : : : : ||: : : : :  
1 MASALSYSKFKSFVILFTPLLLPLVILMPAKVSCCAYVILMAYIWTCTEVIPLATS 60  
QY 61 LFPVILFPLMGIMDASEVCIEYFKDTNILFVGGGLMVAIAVEHNLHRIALQVLIIGVR 120  
DB ||: : : : : ||: : : : : ||: : : : : ||: : : : : ||: : : : :  
61 LMPVLLFPLFQILDSRQVCVQYMKDTNMLFLGGLIVAVAVERNLHRIALRTLLWVGAK 120  
QY 121 PALLLLGFMVITAFLSMISNTATTAMVPVIGHAVLEBOLQSGKKDVEGG-----NNN 172  
DB ||: : : : : ||: : : : : ||: : : : : ||: : : : : ||: : : : :  
121 PARMLGFMGVTTALLSMWISNTATTAMVPVIVEAILQOMEATSAATEAGLEGGQTNNL 180  
QY 173 PTFE-----LQEECPQ-----KEVTKL-----DNGQPVSAAPSEP 201  
DB ||: : : : : ||: : : : : ||: : : : : ||: : : : : ||: : : : :  
181 NAEEDDTKAVILGGKCAVIAISTYVKKVKQLINMLTPLKLEKQEQDQLGPIRQDSA 240  
QY 202 RTQKTQEHHRFSQGLSLCICYSASIGGIATLTGTPNLVLOQVNSLFPQNGNVNFAW 261  
DB ||: : : : : ||: : : : : ||: : : : : ||: : : : : ||: : : : :  
241 QCQEDQERKRLCKAMTLCICYAASIGGTATLTGTGNVLLQGMNELFPDSKDLNFAW 300  
QY 262 FGFAPPTMIILLLLAWLQVLFLGVNPKNPFGEGBEERKQAAFOVIKTOYRLLGPMS 321  
DB ||: : : : : ||: : : : : ||: : : : : ||: : : : : ||: : : : :  
301 FAFAPFNMLVMLFAWLQFVYMFSSFKSKSGCGLESKKNEKAAKVLQEEYRKLGPUS 360  
QY 322 FAEKTVTVLVLLVLFVLTREPFPGFGMDTVFANEKGQSMASDGTVAIFISLVMFIIPS 381  
DB ||: : : : : ||: : : : : ||: : : : : ||: : : : : ||: : : : :  
361 FAEINVLCFFLLVILWFSRDFGFMGLTVAWVEGETKSV-SDATVAIFVATLLFIVPS 419  
QY 382 KIP-----GLMQDPKPKGKLKAPPAILTWKTVNDKMPNIVILLGGGFALAGSEOSGLS 436  
DB ||: : : : : ||: : : : : ||: : : : : ||: : : : : ||: : : : :  
420 QKPKNFRSOTBEGKSP-VLIAAPPPLLDWKVQKVPWGIIVLLGGGFALAGSEASGLS 478  
QY 437 EWLGDKLTPLOHPPSATAVILCLIAIETECTSNVATTTLFLPILASMAOAIQLHPLVY 496  
DB ||: : : : : ||: : : : : ||: : : : : ||: : : : : ||: : : : :  
479 VMNGKMEPLHAPPAAITLILSLVAVFECTSNVATTTLFLPISMSRSIGLNPLVI 538  
QY 497 MLPCTLAASLAFMLPVATPPNAIVFSFGGLKVSMDMARAGFLNIIIGLVAITISNWSIP 556  
DB ||: : : : : ||: : : : : ||: : : : : ||: : : : : ||: : : : :  
539 MLPCTLSASFAPMLPVATPPNAIVFYTYGHLKVDADWYKTVIMNIIIGVFCVFLAVTWGRA 598  
QY 557 IFKLDTPFSPWAH 568  
DB ||: : : : : ||: : : : : ||: : : : : ||: : : : : ||: : : : :  
599 IFDLDFPDWAN 610

Search completed: June 30, 2005, 08:52:06  
Job time : 84.8208 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:36:57 ; Search time 22.5692 Seconds  
(without alignments)  
1941.542 Million cell updates/sec

Title: US-10-017-479A-4

Perfect score: 3056

Sequence: 1 MATCHPALWAYRFLYLCL.....HSNTSQCLLNPSNSTVPGGL 587

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 segs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Issued Patents\_AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2392	78.2	599	4	US-09-949-016-9866 Sequence 9866, Ap
2	1325	43.3	601	4	US-09-949-016-9977 Sequence 9977, Ap
3	1259	41.2	627	4	US-09-556-916-8 Sequence 8, Appli
4	1257	41.1	627	4	US-09-556-916-2 Sequence 2, Appli
5	1256	41.1	627	4	US-09-949-016-6840 Sequence 6840, Ap
6	1255.5	41.1	626	4	US-09-556-916-20 Sequence 20, Appl
7	1254	41.0	627	4	US-09-556-916-10 Sequence 10, Appl
8	1253.5	41.0	626	4	US-09-556-916-14 Sequence 14, Appl
9	1252	40.9	627	4	US-09-556-916-4 Sequence 4, Appli
10	1250.5	40.9	626	4	US-09-556-916-22 Sequence 22, Appl
11	1248.5	40.8	626	4	US-09-556-916-15 Sequence 15, Appl
12	1097	35.9	581	4	US-09-556-916-12 Sequence 12, Appl
13	1095	35.8	581	4	US-09-556-916-6 Sequence 6, Appli
14	1093.5	35.8	580	4	US-09-556-916-24 Sequence 24, Appl
15	1091.5	35.7	580	4	US-09-556-916-18 Sequence 18, Appl
16	1023	35.5	561	4	US-09-949-016-8161 Sequence 8161, Ap
17	776	25.4	527	4	US-09-602-787A-516 Sequence 516, App
18	570.5	18.7	368	4	US-09-270-767-42029 Sequence 42029, A
19	522	17.1	470	4	US-09-543-681A-5952 Sequence 5952, Ap
20	493	16.1	233	4	US-09-489-847-176 Sequence 176, App
21	475	15.5	335	4	US-09-602-787A-518 Sequence 518, App
22	407.5	13.3	524	3	US-09-134-001C-5457 Sequence 5457, Ap
23	314.5	10.3	230	4	US-09-270-767-43713 Sequence 43713, A
24	307.5	10.1	180	4	US-09-270-767-42669 Sequence 42669, A
25	281.5	9.2	194	4	US-09-270-767-59102 Sequence 59102, A
26	279	9.1	421	4	US-09-248-796A-20749 Sequence 20749, A
27	250	8.2	547	4	US-09-489-039A-13843 Sequence 13843, A

28	247	8.1	132	4	US-09-270-767-57987 Sequence 57987, A
29	233	7.6	169	4	US-09-270-767-57286 Sequence 57286, A
30	233	7.6	596	4	US-09-902-540-13547 Sequence 13547, A
31	229	7.5	548	4	US-09-902-540-11870 Sequence 11870, A
32	228	7.5	624	4	US-09-543-681A-4343 Sequence 4343, Ap
33	212	6.9	614	4	US-09-489-039A-12605 Sequence 12605, A
34	210.5	6.9	493	4	US-09-540-236-2120 Sequence 2120, Ap
35	208	6.8	302	4	US-09-902-540-10445 Sequence 10445, A
36	205.5	6.7	694	4	US-09-252-991A-22637 Sequence 22637, A
37	196	6.4	494	4	US-09-543-681A-7033 Sequence 7033, Ap
38	169	5.5	478	3	US-09-134-001C-4637 Sequence 4637, Ap
39	155.5	5.1	449	4	US-09-328-352-7512 Sequence 7512, Ap
40	154	5.0	470	4	US-09-438-185A-209 Sequence 209, App
41	148	4.4	443	4	US-09-602-787A-532 Sequence 532, App
42	138.5	4.5	211	4	US-09-198-452A-1166 Sequence 1166, Ap
43	137.5	4.5	597	4	US-09-489-039A-14112 Sequence 14112, A
44	133	4.3	381	4	US-09-710-279-3284 Sequence 3284, Ap
45	128	4.2	524	4	US-09-198-452A-369 Sequence 369, App
46	128	4.2	524	4	US-09-438-185A-353 Sequence 353, App
47	125.5	4.1	430	3	US-09-134-001C-2981 Sequence 2981, Ap
48	125	4.1	515	4	US-09-869-433-2 Sequence 2, Appli
49	124	4.1	619	4	US-09-540-236-2377 Sequence 2377, Ap
50	120.5	3.9	437	4	US-09-543-681A-6984 Sequence 6984, Ap
51	117.5	3.8	398	4	US-09-489-039A-13159 Sequence 13159, A
52	117	3.8	603	3	US-09-097-889-23 Sequence 23, Appl
53	117	3.8	603	4	US-09-098-079-23 Sequence 23, Appl
54	116.5	3.8	526	4	US-09-543-681A-8223 Sequence 8223, Ap
55	116	3.8	518	4	US-09-252-991A-30478 Sequence 30478, A
56	115.5	3.8	441	4	US-09-489-039A-10612 Sequence 10612, A
57	115.5	3.8	502	4	US-09-252-991A-23327 Sequence 23327, A
58	113.5	3.7	429	4	US-08-311-731A-287 Sequence 287, App
59	113	3.7	363	4	US-09-270-767-42613 Sequence 42613, A
60	111	3.6	482	4	US-09-328-352-7784 Sequence 7784, Ap
61	108.5	3.5	616	4	US-09-489-039A-12599 Sequence 12599, A
62	108	3.5	167	4	US-09-248-796A-20740 Sequence 20740, A
63	108	3.4	319	4	US-09-270-767-57927 Sequence 57927, A
64	105.5	3.4	485	4	US-09-134-000C-6031 Sequence 6031, Ap
65	105.5	3.4	533	1	US-08-294-872-2 Sequence 2, Appli
66	105.5	3.4	533	5	PCT-US95-09823-2 Sequence 2, Appli
67	105	3.4	907	3	US-08-863-102-2 Sequence 2, Appli
68	105	3.4	996	4	US-09-252-991A-27018 Sequence 27018, A
69	103.5	3.4	345	4	US-09-902-540-9836 Sequence 9836, Ap
70	103.5	3.4	529	4	US-09-489-039A-8824 Sequence 8824, Ap
71	103.5	3.4	549	4	US-09-489-039A-13779 Sequence 13779, A
72	103	3.4	459	4	US-09-489-039A-10630 Sequence 10630, A
73	102.5	3.4	729	4	US-09-107-532A-3970 Sequence 3970, Ap
74	102	3.3	116	4	US-09-602-787A-520 Sequence 520, App
75	102	3.3	569	2	US-08-750-723A-2 Sequence 2, Appli
76	102	3.3	569	3	US-09-191-275-2 Sequence 2, Appli
77	101.5	3.3	470	4	US-09-328-352-8240 Sequence 8240, Ap
78	101.5	3.3	638	4	US-09-252-991A-18036 Sequence 18036, A
79	101	3.3	509	4	US-09-252-991A-23341 Sequence 23341, A
80	101	3.3	614	4	US-09-543-681A-4330 Sequence 4330, Ap
81	100.5	3.3	436	4	US-09-583-110-4729 Sequence 4729, Ap
82	100.5	3.3	440	4	US-09-107-433-5201 Sequence 5201, Ap
83	100.5	3.3	463	4	US-09-252-991A-25488 Sequence 25488, A
84	99.5	3.3	456	3	US-09-134-001C-3771 Sequence 3771, Ap
85	99	3.2	898	3	US-08-863-102-1 Sequence 1, Appli
86	99	3.2	898	3	US-08-863-102-4 Sequence 4, Appli
87	98.5	3.2	492	4	US-09-489-039A-8343 Sequence 8343, Ap
88	98.5	3.2	652	3	US-08-987-123-2 Sequence 2, Appli
89	97.5	3.2	402	4	US-09-489-039A-10525 Sequence 10525, A
90	97.5	3.2	413	4	US-09-543-681A-5608 Sequence 5608, Ap
91	97.5	3.2	552	4	US-09-248-796A-20151 Sequence 20151, A
92	97	3.2	641	4	US-09-489-039A-11893 Sequence 11893, A
93	97	3.2	641	4	US-09-543-681A-4453 Sequence 4453, Ap
94	96	3.1	477	4	US-09-252-991A-29825 Sequence 29825, A
95	96	3.1	514	4	US-09-252-991A-25845 Sequence 25845, A
96	95.5	3.1	391	4	US-09-710-279-1236 Sequence 1236, Ap
97	95.5	3.1	422	3	US-09-134-001C-4048 Sequence 4048, Ap
98	95.5	3.1	442	4	US-09-809-665A-73 Sequence 73, Appl
99	95.5	3.1	448	4	US-09-134-000C-3794 Sequence 3794, Ap
100	95.5	3.1	482	4	US-09-328-352-6084 Sequence 6084, Ap

ALIGNMENTS

```

RESULT 1
US-09-949-016-9866
; Sequence 9866, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9866
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9866

Query Match      78.2%; Score 2392; DB 4; Length 599;
Best Local Similarity 76.9%; Pred. No. 6e-235;
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;

QY 1 MATCPALWAYRYLIIVLCPIFLPIVQTKAEYCAYSIIIMALLMCTEALPLAVTA 60
DB 8 MATCQALWAYRSYLIIVFFPIILLPLIPVPSKEAYCAVAIIIMALFCTEALPLAVTA 67

QY 61 LPPIVLPMLGIMDASEVCIEYFKDTNIIIFVGLGMVAIAVEHNLKRIALQVLLIIGVR 120
DB 68 LPPIILFPMMGIVDASEVAYEYKDSNLLFPFGLLMVAIAVEHNLKRIALRVLLIIGVR 127

QY 121 PALLILGMLVTAFLSMWISNTATTAMVPIGHAVLEQLOGSK--KDVEGNNPTPELO 178
DB 128 PAPILGMLVTAFLSMWISNTATAMVPFAHVAVDQLHSSQASSNVEGNNPTPELO 187

QY 179 EECQKEVTKLDNGQ--PV-SAPSEPTQKQEHRRFSQGLSICICYSASIGGIATLTGT 235
DB 188 EPSQKEVTKLDNGQALPVTSSSEGRAHLSQKHLLTQCMSLCVCYSASIGGIATLTGT 247

QY 236 TPNVLVQGVNSLPQNGVNVNFASFQFAFPTMIILLALLAWLQVLFVGNFRKNPGF 295
DB 248 APNLVLOQVNSLPQNGVNVNFASFQFAFPTMIILLALLAWLQVLFVGNFRKNPGI 307

QY 296 GEGEERKQAAFOVTKQYRLGPMPSFAEKTVTVLVLLVLMFTREBPGFPFGMGDTVFA 355
DB 308 CEKQEQQAAYCVIQTHERLLGPMTPFAKSAISILFVLLVLMFTREBPGFPGWNLAFP 367

QY 356 NEQGQSMASDGTVAIFISLVNFIIPSKIPGLMQDPKPKGKUKAPPAIILTKTVNDKMPWN 415
DB 368 NAKGESMVSDGTVAIFIGIIFIIIPSKPPGLTQDPENPGKLPKAPGLLDKMTVQKMPWN 427

QY 416 IVILLGGFALAKSGEQSLSEWLGDKLTPLQHPSPSATVILCILLTAIFTECTSNVATT 475
DB 428 IVLLGGGYALAKSGERSGLSEWLGDKLTPLQHPSPSATVILCILLTAIFTECTSNVATT 487

QY 476 TLFPLILASMAQAICLHPYVYMLPCTLAASLAFMLPVATPPNAIVSFSGGLKVSDMARAG 535
DB 488 TIFLPLILASMAQAICLHPYVYMLPCTLAATLAFMLPVATPPNAIVSFSGGLKVLDMARAG 547

QY 536 FLNLTIGVLATLTSINSWSIPIFKLDTPTPSWAHNSV-TSQCLLNPSNSTVP 584
DB 548 FLNLTIGVLATLTSINSWSIPIFKLDTPTPSWAHNSV-TSQCLLNPSNSTVTP 597

```

```

RESULT 2
US-09-949-016-9977
; Sequence 9977, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9977
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9977

```

```

Query Match      43.3%; Score 1325; DB 4; Length 601;
Best Local Similarity 44.4%; Pred. No. 4.6e-126;
Matches 267; Conservative 125; Mismatches 167; Indels 42; Gaps 11;

QY 11 YRFYILVLCPIFLPIVQTKAEYCAYSIIIMALLMCTEALPLAVTALPVLPLM 70
DB 16 YRFLFVFTVLVLLPLPIVHTKEACAYTLFVVATFWLTEALPLSVTALLPSLMPLMF 75

QY 71 GIMDASEVCIEYFKDTNIIIFVGLGMVAIAVEHNLKRIALQVLLIIGVRPALLIGFML 130
DB 76 GIMPSKVASAYFKDFHLLIGVICLATSIERNLKHRIALQVMMVGVNPNALTLGFS 135

QY 131 VTAFLSMWISNTATTAMVPIGHAVLEQLOGSKQDVE-----GNNNPTPELOEC-- 181
DB 136 STAFLSMWISNTATTAMVPFAHVAVDQLHSSQASSNVEGNNPTPELO 195

QY 182 ----PQKEVTK-----LDNGQ-----PVSAPSEPTQ-KTOEHRRFSQGLSICICYSAS 225
DB 196 HEINERKENTKPVPGYNNDTGKISSKVELEKNSGMEKTKYRTRKGVTRKLTCLCIAYSST 255

QY 226 IGGIATLTGTPNLVLOQGVNSLPQNGVNVNFASFQFAFPTMIILLALLAWLQVLF 285
DB 256 IGGLTITGTSTNLIFAEYFNTRY-PDCRLNFGSMFTSFPAALIIILLSLWILQWLF 314

QY 286 GVNFRKNFGFEGEERKQAAFOVTKQYRLGPMPSFAEKTVTVLVLLVLMFTREBPGF 345
DB 315 GFNFKEFKGKTKTVQKACAEVIAKQKLGPIRYQELVTLVLIIMALLWFSRDPGF 374

QY 346 FPGMGDTVFANEXQSMASDGTVAIFISLVNFIIPSKIPGLMQDPKPKGKKA--PPAIL 403
DB 375 VPCW-SALFSEYFG--FATDSTVALLIGLFLFIPAK---TLTKTTGTGEIVAFDYSPLI 428

QY 404 TWKTVNDKMPWNIVILLGGFALAKSGEQSLSEWLGDKLTPLQHPSPSATVILCILLIA 463
DB 429 TWKEFQSFMPWDIAIIVGGGFALADGCESSGKWNKSLPGSLPAWLIILISSLMVT 488

QY 464 IFTECTSNVATTTLFPLILASMAQAICLHPYVYMLPCTLAASLAFMLPVATPPNAIVSF 523
DB 489 SLTEVASNPATYTLFPLILSPLAEAHVNPYLIILISTLCTSFALLPVANPENAIVSF 548

QY 524 GGLKVSMDARAGFLNLTIGVLATLTSINSWSIPIFKLDTPTPSWAHNSVTSQCLLNPSNSTV 583
DB 549 GHLKVIDMWKAGLVNIGVAVVWMLGICTWIVPMFDLYTPSWAPA-----MSNETM 600

QY 584 P 584

```

Db 601 P 601

## RESULT 3

US-09-556-916-8  
; Sequence 8, Application US/09556916  
; Patent No. 6548271; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Nehls, Michael

; APPLICANT: Friedrich, Glenn

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins

; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916

; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 8

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-556-916-8

Query Match 41.2%; Score 1259; DB 4; Length 627;  
Best Local Similarity 41.2%; Pred. No. 2.7e-119;

Matches 253; Conservative 125; Mismatches 172; Indels 64; Gaps 13;

QY 12 RYLVLCPLPILPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTALPFIIVPLPLMG 71

Db 11 RKLWVCPVLLPLPLPHPSSEACAYLVITAVYVWSEAVPLGAAALVPAFLYPPFG 70

QY 72 IMDASEVCIYEKDTNIIILFVGLMVAIAVEHWNHKLRIALQVLLIIGVRPALLLGFMLV 131

Db 71 VLSNEVAAYEYFNTLLVGVICVAAVEKNLHKLRIALRMVLMAGAKPGMLLCFMCC 130

QY 132 TAFLSMISNTATTAMVPIGHAVLEQLOGSK-KDVEGNNN-----PTF 175

Db 131 TTLLSMLSNTSTTAMVPIVEAVLQELVSADEQLVAGNSNTEEAEPISLDVKNQPSL 190

QY 176 EL---QECPQKEVKLD-----NQP-----VSAPSEPR 202

Db 191 ELIFVNEDRSNADLTTLMHENLNGVPSITNPITKANQHKKQHPQVLTPTS-PR 249

QY 203 TQKT---QEHH--RFSQGLSLCICYSASIGGIATLTGTPNLVLOGVNSLFPQNGNV 256

Db 250 KQKLRKYRSHDDQMIKCLSLISYSATIGLTTIIGTSLIFLEHFNQYPA-AEVV 308

QY 257 NFASWFGFAPFTMIILLALLMQLVFLGVNFRKNFGFGEGBE-ERKQAAPOVKTQYR 315

Db 309 NFGTWFLSFPISLIMLVSWFMHMLFLGNCFKETCSLSKKKTKRQLSEKRIQEBYE 368

QY 316 LIGPMSFAEKTVTVLVLLVLMVFTREPGFPFGWDTVFANEKGOSMASDGTVAIFISLV 375

Db 369 KLGDISYEMVTGFFFIIMTLVMTVLTREPGFVGWDS--FFEKKG--YRTDATVSFVLGFL 424

QY 376 MFIIPSKIP--GLMOPDKPKGLKAPPAITLTKTVDKMPNIVILLGGGALAGSQS 433

Db 425 LFLIPAKKPCFKKNDGENQHSLSGTBIITWKDFQKTPMPEIIVLVGGVALAGSKSS 484

QY 434 GLSEWLGDKLTPLQHIIPPSATAVILCLLIAITFECTSNVATTTFLPLILASMAQAICLHP 493

Db 485 GLSTWIGNQMLSSLPLPWAVTLACILSVISVTEFVSNPATITIFPLILCSLSETLHP 544

QY 494 LYVMLPCTLAASLAPLVPATPPNAIVSFGGLKVDMDARAGFLNITGLVAILTISINSW 553

Db 545 LYTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQIKDMVKAGLVNVIGLVIMVAINTW 604

QY 554 SIPIFKLDTFPSWA 567

Db 605 GVSLPHLDTPAWA 618

Db

RESULT 5

## RESULT 4

US-09-556-916-2

; Sequence 2, Application US/09556916  
; Patent No. 6548271; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Nehls, Michael

; APPLICANT: Friedrich, Glenn

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins

; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916

; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-556-916-2

Query Match 41.1%; Score 1257; DB 4; Length 627;  
Best Local Similarity 41.2%; Pred. No. 4.3e-119;

Matches 253; Conservative 124; Mismatches 173; Indels 64; Gaps 13;

QY 12 RYLVLCPLPILPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTALPFIIVPLPLMG 71

Db 11 RKLWVCPVLLPLPLPHPSSEACAYLVITAVYVWSEAVPLGAAALVPAFLYPPFG 70

QY 72 IMDASEVCIYEKDTNIIILFVGLMVAIAVEHWNHKLRIALQVLLIIGVRPALLLGFMLV 131

Db 71 VLSNEVAAYEYFNTLLVGVICVAAVEKNLHKLRIALRMVLMAGAKPGMLLCFMCC 130

QY 132 TAFLSMISNTATTAMVPIGHAVLEQLOGSK-KDVEGNNN-----PTF 175

Db 131 TTLLSMLSNTSTTAMVPIVEAVLQELVSADEQLVAGNSNTEEAEPISLDVKNQPSL 190

QY 176 EL---QECPQKEVKLD-----NQP-----VSAPSEPR 202

Db 191 ELIFVNEDRSNADLTTLMHENLNGVPSITNPITKANQHKKQHPQVLTPTS-PR 249

QY 203 TQKT---QEHH--RFSQGLSLCICYSASIGGIATLTGTPNLVLOGVNSLFPQNGNV 256

Db 250 KQKLRKYRSHDDQMIKCLSLISYSATIGLTTIIGTSLIFLEHFNQYPA-AEVV 308

QY 257 NFASWFGFAPFTMIILLALLMQLVFLGVNFRKNFGFGEGBE-ERKQAAPOVKTQYR 315

Db 309 NFGTWFLSFPISLIMLVSWFMHMLFLGNCFKETCSLSKKKTKRQLSEKRIQEBYE 368

QY 316 LIGPMSFAEKTVTVLVLLVLMVFTREPGFPFGWDTVFANEKGOSMASDGTVAIFISLV 375

Db 369 KLGDISYEMVTGFFFIIMTLVMTVLTREPGFVGWDS--FFEKKG--YRTDATVSFVLGFL 424

QY 376 MFIIPSKIP--GLMOPDKPKGLKAPPAITLTKTVDKMPNIVILLGGGALAGSQS 433

Db 425 LFLIPAKKPCFKKNDGENQHSLSGTBIITWKDFQKTPMPEIIVLVGGVALAGSKSS 484

QY 434 GLSEWLGDKLTPLQHIIPPSATAVILCLLIAITFECTSNVATTTFLPLILASMAQAICLHP 493

Db 485 GLSTWIGNQMLSSLPLPWAVTLACILSVISVTEFVSNPATITIFPLILCSLSETLHP 544

QY 494 LYVMLPCTLAASLAPLVPATPPNAIVSFGGLKVDMDARAGFLNITGLVAILTISINSW 553

Db 545 LYTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQIKDMVKAGLVNVIGLVIMVAINTW 604

QY 554 SIPIFKLDTFPSWA 567

Db 605 GVSLPHLDTPAWA 618

```
US-09-949-016-6840
; Sequence 6840, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6840
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6840

Query Match          41.1%; Score 1256; DB 4; Length 627;
Best Local Similarity 41.2%; Pred. No. 5.4e-119;
Matches 253; Conservative 124; Mismatches 173; Indels 64; Gaps 13;

QY 12 RFYLIVLCPIELLPLIVQTKYACAYSIILMALLWCTEALPLAVTALFPVLPPLMG 71
DB 11 RKLVLVVCVPLLLPLPLVHPSEASCAYLVITAVYVWSEAVPLGAAALVPAFLYPPFG 70
QY 72 IMDASEVCIFYKDTNTLIFVGGMLVAIAVEHNNLHKRIALQVLLIIGVRPALLLGFMLV 131
DB 71 VLRSNEVAEYFKNTLLLVGVICVAAAEKNNLHKRIALRWLMAGAKPGLLLCFMCC 130
QY 132 TAPLSMWISNTATTAMVPIGHAVLEQLQSGK-KOVEGGNN-----PTF 175
DB 131 TTLLSMWLSNTSTTAMVMPIVEAVLQELVSAEDELVAGNSNTEEAAPISLDVKNQPSL 190
QY 176 EL---QECPQKVTCLDNGQPVSA-----NGQP-----VSAPSEPR 202
DB 191 ELIFVNEEDRSNADTLTMHNENLNGVPSITNPIKTANQHQKQHPQSEKPVLTSPR 249
QY 203 TQKT-----QEHH--RFSGSLCICYSASIGGATLTGTPNLVLOQVNSLPQNGNVV 256
DB 250 KQKLNRYRSHDDQMICKLSLSISYSATIGGLTIIIGTSLIFLHFNNQYPA-AEVV 308
QY 257 NFASWGFAPPTMIILLLLAWLQVLFVGNFRKNFGFGEGBE-ERKQAAFOVIKTOYR 315
DB 309 NFGTWFLSPISLIMLVSWFMWHLFLGCNFKETCSLKKKTKREQLSEKRIQEBYE 368
QY 316 LLGPMSPAETVTVLVLLVLMFTREPFGPPGMDTVFANEKQSMASDGTVAIFISLV 375
DB 369 KJGDISYEMVTGFFILMTVLMFTREPFGVPGWDS--FFEKKG--YRTDATVSVFLGL 424
QY 376 MFIIPSKIP--GLMQDPKPKGLKAPAILTWKTVNDKMPNIVILLGGGFALAKGSEQS 433
DB 425 LFLIPAKKPCFGKNDGENQBSLGTESIITWKDQKTMPEIIVILVGGGYALASGSKS 484
QY 434 GLSWMGDKLTPLQHPISPSATAVILCLLIAIFTECTSNVATTLFLPILASMAQICLHP 493
DB 485 GLSTWIGNQMSLSLPPWAVTLIACILSVISVTEFVSNPATITIFLILCSLSETMHNIP 544
QY 494 LYVNLPTCLAAASLAPMLPVATPPNAIVFSFGGLKVSMDARAGFLNIIIGVLAITLSINSW 553
DB 545 LYTLLIPVTMCISFAVMLPVGNPPNAIVFSYGHCOIKDMVKAGLVNIGLVIMVAINTW 604
QY 554 SIPIEKLDTPPSWA 567
DB 605 GVSLFHLDTYPAWA 618
```

```
RESULT 6
US-09-556-916-20
; Sequence 20, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NO. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-20

Query Match          41.1%; Score 1255.5; DB 4; Length 626;
Best Local Similarity 40.8%; Pred. No. 6e-119;
Matches 250; Conservative 123; Mismatches 178; Indels 61; Gaps 11;

QY 12 RFYLIVLCPIELLPLIVQTKYACAYSIILMALLWCTEALPLAVTALFPVLPPLMG 71
DB 11 RKLVLVVCVPLLLPLPLVHPSEASCAYLVITAVYVWSEAVPLGAAALVPAFLYPPFG 70
QY 72 IMDASEVCIFYKDTNTLIFVGGMLVAIAVEHNNLHKRIALQVLLIIGVRPALLLGFMLV 131
DB 71 VLRSNEVAEYFKNTLLLVGVICVAAAEKNNLHKRIALRWLMAGAKPGLLLCFMCC 130
QY 132 TAPLSMWISNTATTAMVPIGHAVLEQLQSGK-KOVEGGNN-----PTF 175
DB 131 TTLLSMWLSNTSTTAMVMPIVEAVLQELVSAEDELVAGNSNTEEAAPISLDVKNQPSL 190
QY 176 EL---QECPQKVTCLDNGQPVSA-----NGQP-----VSAPSEPR 202
DB 191 ELIFVNEEDRSNADTLTMHNENLNGVPSITNPIKTANQHQKQHPQSEKPVLTSPR 250
QY 203 TQKT-----QEHH--RFSGSLCICYSASIGGATLTGTPNLVLOQVNSLPQNGNVV 256
DB 251 KLNRYRSHDDQMICKLSLSISYSATIGGLTIIIGTSLIFLHFNNQYPA-AEVV 309
QY 259 ASWGFAPPTMIILLLLAWLQVLFVGNFRKNFGFGEGBE-ERKQAAFOVIKTOYR 317
DB 310 GTWFLSPISLIMLVSWFMWHLFLGCNFKETCSLKKKTKREQLSEKRIQEBYE 369
QY 318 GPMSPAETVTVLVLLVLMFTREPFGPPGMDTVFANEKQSMASDGTVAIFISLV 377
DB 370 GDISYEMVTGFFILMTVLMFTREPFGVPGWDS--FFEKKG--YRTDATVSVFLGL 425
QY 378 IIPSKIP--GLMQDPKPKGLKAPAILTWKTVNDKMPNIVILLGGGFALAKGSEQS 435
DB 426 LIPAKKPCFGKNDGENQBSLGTESIITWKDQKTMPEIIVILVGGGYALASGSKS 485
QY 436 SEWLGDKLTPLQHPISPSATAVILCLLIAIFTECTSNVATTLFLPILASMAQICLHP 495
DB 486 STWIGNQMSLSLPPWAVTLIACILSVISVTEFVSNPATITIFLILCSLSETMHNIP 545
QY 496 VMLPCTCLAAASLAPMLPVATPPNAIVFSFGGLKVSMDARAGFLNIIIGVLAITLSINSW 555
DB 546 TLLIPVTMCISFAVMLPVGNPPNAIVFSYGHCOIKDMVKAGLVNIGLVIMVAINTW 605
QY 556 SIPIEKLDTPPSWA 567
DB 606 SLFHLDTYPAWA 617

RESULT 7
US-09-556-916-10
```

; Sequence 10, Application US/09556916  
; Patent No. 6548271

## ; GENERAL INFORMATION:

; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins

; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916

; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 10

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-556-916-10

Query Match 41.0%; Score 1254; DB 4; Length 627;

Best Local Similarity 41.2%; Pred. No. 8.6e-119;

Matches 253; Conservative 124; Mismatches 173; Indels 64; Gaps 13;

QY 12 RYLVLCPLPFLPLPLIVQKEAYCAYSILMALLMCTEALPLAVTALPVLFPPLMG 71  
DB 11 RKLWVCPVLLPLPLVHPSEASCAVLTAVYVWSEAVPLGAAALVPAFLYPPFG 70  
QY 72 IMDASEVCIEYFKDNTILFVGLMVAIAVEHNLHKLRIALQVLLIIGVRPALLLIGFMLV 131  
DB 71 VLRSNEVAAYEYFKNTLLVGVICVAAAVEKNLHKLRIALRMVLMAKAPGMLLFCFCC 130  
QY 132 TAFLSMISNTATTAMVPIGHAIVLEQLQSGK-KDVEGNNN-----PTF 175  
DB 131 TTLSSMWSNTSTTAMVPIEAVLQELVSADEQLVAGNSNTEAEPISLDVKNQPSL 190  
QY 176 EL---QECPQKEVTKLD-----NQOP-----VSAPSEPR 202  
DB 191 ELIFVNEEDRSNADLTTLMHENENLVGPSITNPITKANQHQKQHPQSEKQVLTFS-PR 249  
QY 203 TQKT----QEHH--RFSQGLSLCICYSASIGGIATLTGTPNLVLOGQVNSLFPQNGVV 256  
DB 250 KQKLRKYSRHHQDMICKLSLSISYSATIGGLTIIIGTSLIFLEHFNQYPA-AEVV 308  
QY 257 NFASWFGFAFTPMIILLALLAWLQVLFVGNFRKNFGEGEE-ERKQAAFOVITQYR 315  
DB 309 NFGTWLFSPISLIMLVSWFMMWLFGLCNFKETCSLSKKKTKRQLSEKRIQEBYE 368  
QY 316 LLGPMSPAEKTVTVLVLLVLMFTREPFGFGMDTVFANEKGQSMASDGTVAIFISLV 375  
DB 369 KLGDISYEMWVGTFPILMTVLMFTREPFGVPGWDS--FPEKKG--YRTDATVSVFLGFL 424  
QY 376 MFIIPSKIP--GLMQDPKPKGLKAPPAILTWKTVNDKMPNIVILLGGGFALAKGSEQS 433  
DB 425 LFLIPAKKCFCKKNDGQNEHSLGTEIITWKDFQKTPWEIVLVGGYALASGSKSS 484  
QY 434 GLSEWLGKLPLOQHPISATVILCLLIAIIFTECTSNVATTLPLPILASMAQAICLHP 493  
DB 485 GLSTWIGNQMLSSLPSPVAVTLLACILSVIVTEFVSNPATIITPILCLSETLHINPLY 544  
QY 494 LYVMLPCTLAASLAFMLPVATPPNAIVFSFGGLKVSMDARAGFLNIIIGVLAITLSINSW 553  
DB 545 LYTLPVTWCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNVLGLVIMVAINTW 604  
QY 554 SIPIKLDTFPSWA 567  
DB 605 GVSFLHLDTPAWA 618

## RESULT 8

US-09-556-916-14

; Sequence 14, Application US/09556916

; Patent No. 6548271

## ; GENERAL INFORMATION:

; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins

; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916

; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 14

; LENGTH: 626

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-556-916-14

Query Match 41.0%; Score 1253.5; DB 4; Length 626;

Best Local Similarity 40.8%; Pred. No. 9.7e-119;

Matches 250; Conservative 122; Mismatches 179; Indels 61; Gaps 11;

QY 12 RYLVLCPLPFLPLPLIVQKEAYCAYSILMALLMCTEALPLAVTALPVLFPPLMG 71  
DB 11 RKLWVCPVLLPLPLVHPSEASCAVLTAVYVWSEAVPLGAAALVPAFLYPPFG 70  
QY 72 IMDASEVCIEYFKDNTILFVGLMVAIAVEHNLHKLRIALQVLLIIGVRPALLLIGFMLV 131  
DB 71 VLRSNEVAAYEYFKNTLLVGVICVAAAVEKNLHKLRIALRMVLMAKAPGMLLFCFCC 130  
QY 132 TAFLSMISNTATTAMVPIGHAIVLEQLQSGK-KDVEGNNN-----PTF 175  
DB 131 TTLSSMWSNTSTTAMVPIEAVLQELVSADEQLVAGNSNTEAEPISLDVKNQPSL 190  
QY 176 EL---QECPQKEVTKLDNGQVSPAS-----EPRTO 204  
DB 191 ELIFVNEESNADLTTLMHENENLVGPSITNPITKANQHQKQHPQSEKQVLTFSPRKQ 250  
QY 205 KT----QEHH--RFSQGLSLCICYSASIGGIATLTGTPNLVLOGQVNSLFPQNGVVNF 258  
DB 251 KLARKYSRHHQDMICKLSLSISYSATIGGLTIIIGTSLIFLEHFNQYPA-AEVVNF 309  
QY 259 ASWFGFAFTPMIILLALLAWLQVLFVGNFRKNFGEGEE-ERKQAAFOVITQYRLL 317  
DB 310 GTWELFSPISLIMLVSWFMMWLFGLCNFKETCSLSKKKTKRQLSEKRIQEBYEKL 369  
QY 318 GPMSPAEKTVTVLVLLVLMFTREPFGFGMDTVFANEKGQSMASDGTVAIFISLVNF 377  
DB 370 GDISYEMWVGTFPILMTVLMFTREPFGVPGWDS--FPEKKG--YRTDATVSVFLGFL 425  
QY 378 IIPSKIP--GLMQDPKPKGLKAPPAILTWKTVNDKMPNIVILLGGGFALAKGSEQSGL 435  
DB 426 LIPAKKCFCKKNDGQNEHSLGTEIITWKDFQKTPWEIVLVGGYALASGSKSSGL 485  
QY 436 SEWLGKLPLOQHPISATVILCLLIAIIFTECTSNVATTLPLPILASMAQAICLHP 495  
DB 486 STWIGNQMLSSLPSPVAVTLLACILSVIVTEFVSNPATIITPILCLSETLHINPLY 545  
QY 496 VMLPCTLAASLAFMLPVATPPNAIVFSFGGLKVSMDARAGFLNIIIGVLAITLSINSWSI 555  
DB 546 TLIPVTWCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNVLGLVIMVAINTW 605  
QY 556 PIPIKLDTFPSWA 567  
DB 606 SLFHLDTYPAWA 617

## RESULT 9

US-09-556-916-4

; Sequence 4, Application US/09556916

; Patent No. 6548271

; GENERAL INFORMATION:

; APPLICANT: Turner, Alex

; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-556-916-4

Query Match 40.9%; Score 1252; DB 4; Length 627;  
Best Local Similarity 41.2%; Pred. No. 1.4e-118;  
Matches 253; Conservative 123; Mismatches 174; Indels 64; Gaps 13;  
QY 12 RFVLIVLCPIFLPLPLIYVOTKEAYCAYSIILMALLMCTEALPLAVTALFPVLPFLMG 71  
Db 11 RKLLEVCVCPILLPLPLVLPSPSEASCAVLIIVTAVYVWSEAVPLGAAALVPAFLYPPFG 70  
QY 72 IMDASEVCIEYKDTNILFVGGMLVAIVAEHNLHRIALQVLLIIGVRPALLLGFMLV 131  
Db 71 VLRSNEVAAYEYKNTTLLVGVICVAAAEKWNLHRIALRMVLMAGAKPGMLLCFMCC 130  
QY 132 TAFLSMWISNTATTAMVPVPIGHVLEQLQSGK-KDVEGGNNN-----PTF 175  
Db 131 TTLLSMWLSNTTAMVMPVIEAVLQELVSAEDEQLVAGNSNTEEAEPISLDVKNQSPSL 190  
QY 176 EL---QECPQKEVTKLD-----NQDP-----VSAPSEPR 202  
Db 191 ELIFVNEEDRNADLTILMHNNENLVPSIINPIKTANQHQKQHPQSEKQPQVLTFS-PR 249  
QY 203 TQKT-----QEHH--RFSQGLSLCICYSASIGGIATLTGTPNVLQGVNSLFPQNGNVV 256  
Db 250 KQKLNRKYRSHHDQMIKCLSLSLSYSATIGGLTIIIGTSTLSIFLEHFNQYPA-AEVV 308  
QY 257 NFASWFGFAPPTMIILLLLAWLQVLFVGNFRKQFGFGEGBE-ERKQAAFOVIKTYR 315  
Db 309 NFGTWFLFSPISLIMLVSWFWHMLFLGCFNKETCSLSKKKKTKREQLSEKRIQEEYE 368  
QY 316 LGPMSFAEKTVTVLVLLVVLVFTREPGFPFGMDTVFANEKQSGMASDGTVAIFISLV 375  
Db 369 KLGDISPEVMVTGFFILMTVLVFTREPGFVPGWDS--FPEKKG--YRTDATVSVFLGFL 424  
QY 376 MFIPSKIP--GLMQDPKPKGKLKAPPAILTWKTVNDKMPWNIIVILGGGFALAKGSEQS 433  
Db 425 LFLIPAKKPCFGKNDGENQEHSLGTEPIITWKDFQKTPMPEIIVLVGGGYALASGSKS 484  
QY 434 GLSEWGDKLTPLQHPSPSATVILCLLIAIFTECTSNVATTLTFLPILASMAQAICLHP 493  
Db 485 GLSTWIGNQMSLSLSPWAVTLACILSVITFVSNPATITIFLPILCSLSETOHINPLY 544  
QY 494 LYVLPCTLAASLAFMLPVATPNAIVFSGGLKVSDMARAGFLNLIIGVLAITLSINSW 553  
Db 545 LYLIPVTMCISFAVMLPVGNPNNAIVFSYGHCOIKDMVKAGLVNIGLVIVMVAINTW 604  
QY 554 SIPIFKLDTFPPSWA 567  
Db 605 GVSIFHLDTYPAWA 618

RESULT 10  
US-09-556-916-22  
; Sequence 22, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael

; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 22  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-556-916-22

Query Match 40.9%; Score 1250.5; DB 4; Length 626;  
Best Local Similarity 40.8%; Pred. No. 2e-118;  
Matches 250; Conservative 122; Mismatches 179; Indels 61; Gaps 11;  
QY 12 RFVLIVLCPIFLPLPLIYVOTKEAYCAYSIILMALLMCTEALPLAVTALFPVLPFLMG 71  
Db 11 RKLLEVCVCPILLPLPLVLPSPSEASCAVLIIVTAVYVWSEAVPLGAAALVPAFLYPPFG 70  
QY 72 IMDASEVCIEYKDTNILFVGGMLVAIVAEHNLHRIALQVLLIIGVRPALLLGFMLV 131  
Db 71 VLRSNEVAAYEYKNTTLLVGVICVAAAEKWNLHRIALRMVLMAGAKPGMLLCFMCC 130  
QY 132 TAFLSMWISNTATTAMVPVPIGHVLEQLQSGK-KDVEGGNNN-----PTF 175  
Db 131 TTLLSMWLSNTTAMVMPVIEAVLQELVSAEDEQLVAGNSNTEEAEPISLDVKNQSPSL 190  
QY 176 EL---QECPQKEVTKLDNGQPVSAAPS-----EPRTQ 204  
Db 191 ELIFVNEESNADLTILMHNNENLVPSIINPIKTANQHQKQHPQSEKQPQVLTFSPRQ 250  
QY 205 KT-----QEHH--RFSQGLSLCICYSASIGGIATLTGTPNVLQGVNSLFPQNGNVNP 258  
Db 251 KLNRYRSHHDQMIKCLSLSLSYSATIGGLTIIIGTSTLSIFLEHFNQYPA-AEVVNP 309  
QY 259 ASWFGFAPPTMIILLLLAWLQVLFVGNFRKQFGFGEGBE-ERKQAAFOVIKTYRLL 317  
Db 310 GTWFLFSPISLIMLVSWFWHMLFLGCFNKETCSLSKKKKTKREQLSEKRIQEEYEKL 369  
QY 318 GPMSFAEKTVTVLVLLVVLVFTREPGFPFGMDTVFANEKQSGMASDGTVAIFISLVNP 377  
Db 370 GDISPEVMVTGFFILMTVLVFTREPGFVPGWDS--FPEKKG--YRTDATVSVFLGFLF 425  
QY 378 IPTSKIP--GLMQDPKPKGKLKAPPAILTWKTVNDKMPWNIIVILGGGFALAKGSEQS 435  
Db 426 LIPAKKPCFGKNDGENQEHSLGTEPIITWKDFQKTPMPEIIVLVGGGYALASGSKSGL 485  
QY 436 SEWLGDKLTPLQHPSPSATVILCLLIAIFTECTSNVATTLTFLPILASMAQAICLHP 495  
Db 486 STWIGNQMSLSLSPWAVTLACILSVITFVSNPATITIFLPILCSLSETOHINPLY 545  
QY 496 VMLPCTLAASLAFMLPVATPNAIVFSGGLKVSDMARAGFLNLIIGVLAITLSINSW 555  
Db 546 TLIPVTMCISFAVMLPVGNPNNAIVFSYGHCOIKDMVKAGLVNIGLVIVMVAINTW 605  
QY 556 PIPIFKLDTFPPSWA 567  
Db 606 SLFHLDTYPAWA 617

RESULT 11  
US-09-556-916-16  
; Sequence 16, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-556-916-16

Query Match 40.8%; Score 1248.5; DB 4; Length 626;  
Best Local Similarity 40.8%; Pred. No. 3.1e-118;  
Matches 250; Conservative 121; Mismatches 180; Indels 61; Gaps 11;  
QY 12 RFYLIVLCPIFLPLPLIVQTEAYCAYSITLMALLMCTEALPLAVTALFPIVPLPLMG 71  
DB 11 RKLLVVCVPLLLPLVLPHPSEASCAYLVITAVYVSEAVPLGAAALVPAFLYPPFG 70  
QY 72 INDASEVCIEYFKDNTNLPVGGMLVAIAVEHNLHRIALQVLLIIGVRPALLLGLFMLV 131  
DB 71 VLRSNEVAAYEYFQNTLLLVGVICVAAAEKWNHRIALRMVLMAGAKPGMLLLCFMCC 130  
QY 132 TAFLSMWISNTATTAMVPVIGHAVLEQLOGSK-KDVEGGNN-----PTF 175  
DB 131 TTLSSMWLSNTSTTAMVPIVEAVLQELVSAEDEQLVAGNSNTEEAEPISLDVKNQSPSL 190  
QY 176 EL---QECPQKEVTKLDNGQPVSAAPS-----EPRQ 204  
DB 191 ELIFVNEESNADLTLMHNNENLVPSITNPIKTANQHGKQKHPQVLTSPRKQ 250  
QY 205 KT-----QEHH--RFSQGLSLCICYSASIGGIATLTGTPNVLQGVNSLFPQNGVNF 258  
DB 251 KLNRYRSHDDQWICKLSLSISYSATIGLTIIGTSTLSIFLEHFNQYPA-AEVNF 309  
QY 259 ASWFGFAPFTMILLALLWLQVFLGVNFRKNFGFGEGB-ERKQAAFOVKTQYRL 317  
DB 310 GTWFLFSPISLIMLVSWFMHMLFLGCFNFKETCSLSKKKTKREQLSEKIQEYK 369  
QY 318 GPMSPAEKTVTVLVLVLLVMTREBPGFPGGDTVPANEXQSMASDGTVAIFISLWPF 377  
DB 370 GDISPEVMVTGFFILMTVLTREBPGFPGGDS--FPEKKG--YRTDATVSFVLGLF 425  
QY 378 IIPSKIP--GLMQDPKPKGLKAPPAILTWTNDKMPNIVILLGGGFALAKSGEQL 435  
DB 426 LIPAKKPCFGKNDGENQHSGLTETPIITWKDFQKTPMPEIIVLGGGVALASGSKSGL 485  
QY 436 SEWLGDKLTPLQHIIPPSATAVILCLLIAIFTECTSNVATTTLPILASMAQICLHP 495  
DB 486 STWIGNQMLSLSLPPEWAVTLLACILSVITFEVSNPATITITFLICLSLSETHINPLY 545  
QY 496 VMLPCTLAASLAPMLPVATPPNAIVPSFGLKVSMDARAGFLNIIGVLAITLSINSWI 555  
DB 546 TLIPVTMCISFAMLPVGNPPNAIVPSYGHQIKDMVKAGLVNIGLVIMVAINTWGV 605  
QY 556 PIFKLDTPPSWA 567  
DB 606 SLFHLDTYPAWA 617

RESULT 12  
US-09-556-916-12  
; Sequence 12, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT FILING DATE: 2000-04-21

; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 12  
; LENGTH: 581  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-556-916-12

Query Match 35.9%; Score 1097; DB 4; Length 581;  
Best Local Similarity 37.6%; Pred. No. 8e-103;  
Matches 231; Conservative 114; Mismatches 159; Indels 110; Gaps 14;  
QY 12 RFYLIVLCPIFLPLPLIVQTEAYCAYSITLMALLMCTEALPLAVTALFPIVPLPLMG 71  
DB 11 RKLLVVCVPLLLPLVLPHPSEASCAYLVITAVYVSEAVPLGAAALVPAFLYPPFG 70  
QY 72 INDASEVCIEYFKDNTNLPVGGMLVAIAVEHNLHRIALQVLLIIGVRPALLLGLFMLV 131  
DB 71 VLRSNEVAAYEYFQNTLLLVGVICVAAAEKWNHRIALRMVLMAGAKPGMLLLCFMCC 130  
QY 132 TAFLSMWISNTATTAMVPVIGHAVLEQLOGSK-KDVEGGNN-----PTF 175  
DB 131 TTLSSMWLSNTSTTAMVPIVEAVLQELVSAEDEQLVAGNSNTEEAEPISLDVKNQSPSL 190  
QY 176 EL---QECPQKEVTKLDNGQPVSAAPS-----VSAPSEPR 202  
DB 191 ELIFVNEESNADLTLMHNNENLVPSITNPIKTANQHGKQKHPQVLTSPRKQ 249  
QY 203 TQKT-----QEHH--RFSQGLSLCICYSASIGGIATLTGTPNVLQGVNSLFPQNGVNV 256  
DB 250 KQKLNRYRSHDDQWICKLSLSISYSATIGLTIIGTSTLSIFLEHFNQYPA-AEVN 308  
QY 257 NPASWFGFAPFTMILLALLWLQVFLGVNFRKNFGFGEGB-ERKQAAFOVKTQYR 315  
DB 309 NFGTWFLFSPISLIMLVSWFMHMLFLGCFNFKETCSLSKKKTKREQLSEKIQEYK 368  
QY 316 LLGPMSPAEKTVTVLVLVLLVMTREBPGFPGGDTVPANEXQSMASDGTVAIFISLV 375  
DB 369 KLGDISPEVMVTGFFILMTVLTREBPGFPGGDS--FPEKKG--YRTDATVSFVLGLF 424  
QY 376 MFIIPSKIP--GLMQDPKPKGLKAPPAILTWTNDKMPNIVILLGGGFALAKSGEQL 433  
DB 425 LFLIPAKKPCFGKNDGENQHSGLTETPIITWKDFQKTPMPEIIVLGGGVALASGSKS 484  
QY 434 GLSEWLGDKLTPLQHIIPPSATAVILCLLIAIFTECTSNVATTTLPILASMAQICLHP 493  
DB 485 GLSTWIGNQMLSLSLPPEWAVTLLACILSVITFEVSNPATITITFLICLSLSETHINPLY 535  
QY 494 LYVMLPCTLAASLAPMLPVATPPNAIVPSFGLKVSMDARAGFLNIIGVLAITLSINSWI 553  
DB 536 VMLPCTLAASLAPMLPVATPPNAIVPSFGLKVSMDARAGFLNIIGVLAITLSINSWI 558  
QY 554 SIPIFKLDTPPSWA 567  
DB 559 GVSFLHLDTPAWA 572

RESULT 13  
US-09-556-916-6  
; Sequence 6, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21

```

; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-6

```

```

Query Match      35.8%; Score 1095; DB 4; Length 581;
Best Local Similarity 37.6%; Pred. No. 1.3e-102;
Matches 231; Conservative 113; Mismatches 160; Indels 110; Gaps 14;

```

```

QY 12 RYLVLCPIELPLPLIVQTEKAYCAYSIIMALLMCTEALPLAVTALFPVLPPLMG 71
DB 11 RKLIVVCVPLLLPLPLVHPSEASCAVLIIVTAVTWSEAVPLGNAALVPFAFLPFFG 70
QY 72 IMDASEVCIEYKDTNIFVGGMLVAIAVEHNLHRIALQVLLIIGVRPALLLGLFMLV 131
DB 71 VLSRNEVAEYFNKNTLLGVGICVAAVEKNLHRIALRMVLMAGAKPGMLLCFMCC 130
QY 132 TAFSLMWSINTATTAMVPIGHAVLEQLQSGK-KOVEGGNNN-----PTP 175
DB 131 TTLSMWSLNTSTTAMVMPIVEAVLQELVSAEDEQLVAGNSNTEEAEPISLDVKNQPSL 190
QY 176 EL---QECPQKEVTKLD-----NGQP-----VSAPSEPR 202
DB 191 ELIFVNEEDRSNADLTTLHNNENLVPSITNPIKTANQHQKQHPQSEKPVLTFS-PR 249
QY 203 TQKT-----QEHH--RFSQGLSLCICYSASIGGIATLTGTTPNLVLOQGVNSLPQNGNVV 256
DB 250 KQLNRKYRSHDDQMICKLSLSISYSATIGGLTTIIGTSLIFLHFNNQYPA-AEVV 308
QY 257 NFASWGFAPPTMIILLLLAWLQVLFVGNFRKNFGEGEE-ERKQAAFOVIKTOYR 315
DB 309 NFGTWFLSPFISLIMLVSNFWMHFLGCGNFKETCSLKKKTKRQSEKRIQOEYE 368
QY 316 LLGPMSPAETVTVLVLLVLTFTREPGFPFGMDTVFANEKGQSMASDGTVAIFSLV 375
DB 369 KLGDISPEWVTGFFILMTVMTFTREPGFVPGWDS--FPEKKG--YRTDATVSVFLGFL 424
QY 376 MFIIIPSKIP--GLMDPKPKGKAPPAITLTKTNDKMPNIVILGGGFALAKSGEQS 433
DB 425 LFLIPAKKPCFGKKNKGNGENQHSLSGTESIITWKDFQKTPWEIIVLVGGYALASGSKS 484
QY 434 GLSEWGLDKLTPLOHIPPSTAVILCLLIAIFTECTSNVATTLTFLPLASMAQAICLHP 493
DB 485 GLSTWIGNQMLSLSLPPWAVTLACILVSVTFEVSNPATITIFLPILCS-----535
QY 494 LYVMLPCTLAASLAFMPLPVATPPNAIVFSFGLKVSMDARAGFLNIIIGVLAITLSINSW 553
DB 536 -----LVKAGLVNVIGLVIMVVAINTW 558
QY 554 SIPIFKLDTFPPSWA 567
DB 559 GVSFLHLDTPAWA 572

```

```

RESULT 14
US-09-556-916-24
; Sequence 24, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0

```

```

; SEQ ID NO 24
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-24

```

```

Query Match      35.8%; Score 1093.5; DB 4; Length 580;
Best Local Similarity 37.3%; Pred. No. 1.8e-102;
Matches 228; Conservative 112; Mismatches 165; Indels 107; Gaps 12;

```

```

QY 12 RYLVLCPIELPLPLIVQTEKAYCAYSIIMALLMCTEALPLAVTALFPVLPPLMG 71
DB 11 RKLIVVCVPLLLPLPLVHPSEASCAVLIIVTAVTWSEAVPLGNAALVPFAFLPFFG 70
QY 72 IMDASEVCIEYKDTNIFVGGMLVAIAVEHNLHRIALQVLLIIGVRPALLLGLFMLV 131
DB 71 VLSRNEVAEYFNKNTLLGVGICVAAVEKNLHRIALRMVLMAGAKPGMLLCFMCC 130
QY 132 TAFSLMWSINTATTAMVPIGHAVLEQLQSGK-KOVEGGNNN-----PTP 175
DB 131 TTLSMWSLNTSTTAMVMPIVEAVLQELVSAEDEQLVAGNSNTEEAEPISLDVKNQPSL 190
QY 176 EL---QECPQKEVTKLDNGQPVSAPS-----EPRTO 204
DB 191 ELIFVNEESNADLTTLHNNENLVPSITNPIKTANQHQKQHPQSEKPVLTSPRKQ 250
QY 205 KT-----QEHH--RFSQGLSLCICYSASIGGIATLTGTTPNLVLOQGVNSLPQNGNVVNF 258
DB 251 KQLNRKYRSHDDQMICKLSLSISYSATIGGLTTIIGTSLIFLHFNNQYPA-AEVVNF 309
QY 259 ASWFGFAPPTMIILLLLAWLQVLFVGNFRKNFGEGEE-ERKQAAFOVIKTOYRL 317
DB 310 GTWFLSPFISLIMLVSNFWMHFLGCGNFKETCSLKKKTKRQSEKRIQOEYEKL 369
QY 318 GPMSPAETVTVLVLLVLTFTREPGFPFGMDTVFANEKGQSMASDGTVAIFSLVNF 377
DB 370 GDISPEWVTGFFILMTVMTFTREPGFVPGWDS--FPEKKG--YRTDATVSVFLGFLF 425
QY 378 IIPSKIP--GLMDPKPKGKAPPAITLTKTNDKMPNIVILGGGFALAKSGEQSGL 435
DB 426 LIPAKKPCFGKKNKGNGENQHSLSGTESIITWKDFQKTPWEIIVLVGGYALASGSKSGL 485
QY 436 SEWGLDKLTPLOHIPPSTAVILCLLIAIFTECTSNVATTLTFLPLASMAQAICLHP 495
DB 486 STWIGNQMLSLSLPPWAVTLACILVSVTFEVSNPATITIFLPILCS-----534
QY 496 VMLPCTLAASLAFMPLPVATPPNAIVFSFGLKVSMDARAGFLNIIIGVLAITLSINSWSI 555
DB 535 -----LVKAGLVNVIGLVIMVVAINTWGV 559
QY 556 PIPIFKLDTFPPSWA 567
DB 560 SLFHLDTYPAWA 571

```

```

RESULT 15
US-09-556-916-18
; Sequence 18, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 580

```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-18

Query Match      35.7%; Score 1091.5; DB 4; Length 580;
Best Local Similarity 37.3%; Pred. No. 2.9e-102;
Matches 228; Conservative 111; Mismatches 166; Indels 107; Gaps 12;

QY 12 RFYLIIVLCPLPFIPLPLIVQTKYACAYSIIMALLMCTEALPLAVTALFPVLFPPLMG 71
Db 11 RKLLLVVCPVLLPLPLVLPHPSEASCAYLIIVTAVYVSEAVPLGAAALVDAFLYPPFG 70

QY 72 INDASEVCIEYKQTNILFVGGIMVAIAVEHWNLHKRIALQVLLIIGVBPALLILGFMV 131
Db 71 VLRSNEAABEYFQNTLLVGVICVAAAVEKNLHKRIALRMVLMAGAKPGMLLLCFMCC 130

QY 132 TAPLSMWISNTATTAMVPIGHAVLQLOGSK-KOVEGGNNN-----PTF 175
Db 131 TTLLSMWLSNTSTTAMVPIVEAVLQELVSADEQLVAGNSNTEEAEPISLDVKNSQPSL 190

QY 176 EL---QEECPQKEVTKLDNGQPVSAPS-----EPRTQ 204
Db 191 ELIFVNEESNADLTTLHMNENLNGVPSITNPIKTANHQKQHPQOEKQPVLTSPRKQ 250

QY 205 KT-----QEHK--RFSQGLSLCICYSASIGGIATLTGTTNVLVQQQVNSLFPQNGNVNF 258
Db 251 KLNRYRSHDQWICKLSLSYSATIGGLTTIIGTSLIFLHFNNQYPA-AEVVNF 309

QY 259 ASWFGFAPFTMILLLMLQLVLPGLVGNFRKNFGFGGEB-ERKQAAFOVIKTQYRL 317
Db 310 GTWFLFSFISLMLVVSFWFHHWLFGLCNFKETCSLSKKKTKRQJSEKRIQBEYEKL 369

QY 318 GPMSEAKTIVTVLVLLVVLWFTREPPFPFGMDTVFANEKGQSMASDGTVAIFISLVNF 377
Db 370 GDISYEMVTGFFILMTVLMFTREPGFVPGWDS--FPEKKG--YETDATSVFLGFLF 425

QY 378 IIPSKIP--GLMQDPKPKGLKAPPAILTWTVTNDKMPNIVILGGFPALAKGSEQSGL 435
Db 426 LIPAKKPCFCGKNDGENQSHSLGTBPITWKFQKTPWEIVILVGGGYALASGSKSGL 485

QY 436 SEMLGDKLTPLQHIPPSATAVILCLLJALFTECTSNVATTLPLPLASMAQAIChPLY 495
Db 486 STWIGNQMLSSLSPWATVLLACILVSIIVTFVSNPATITIFLPILCS----- 534

QY 496 VMLPCTLAASLAFMLPVATPPNAIVFSGCLKVSDMARAGFLINIIGVLAITLSINWSI 555
Db 535 -----LVKAGLGNNVIGLVVNVVAINTWGV 559

QY 556 PIFKLDTPPSWA 567
Db 560 SLFHLDTYPWA 571
```

Search completed: June 30, 2005, 09:02:16  
Job time : 25.5692 secs

**THIS PAGE IS BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - protein search, using sw model

Run on: June 30, 2005, 08:59:01 ; Search time 76.171 Seconds  
(without alignments)  
2973.073 Million cell updates/sec

Title: US-10-017-479A-4

Perfect score: 3058

Sequence: 1 MATCHPALWAYRYLIVLCL.....HSNTSCLLNPNSTVPGGL 587

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1722976 seqs, 385795295 residues

Total number of hits satisfying chosen parameters: 1722976

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US05\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US04\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US03\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US02\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US01\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US00\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US05\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US04\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US03\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US02\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US01\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US00\_PUBCOMB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_PUBCOMB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US12\_PUBCOMB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US13\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3058	100.0	587	13	US-10-173-519-5
2	3058	100.0	587	14	US-10-167-994-12
3	3058	100.0	587	17	US-10-178-359-13
4	2392	78.2	592	14	US-10-167-994-3
5	2392	78.2	592	14	US-10-167-994-11
6	1971.5	64.5	619	9	US-09-729-094-4
7	1971.5	64.5	619	14	US-10-435-631-4
8	1829	59.8	581	17	US-10-178-359-12
9	1828	53.2	616	15	US-10-092-900A-46
10	1628	53.2	616	15	US-10-092-900A-48
11	1628	53.2	616	15	US-10-403-161-68
					Sequence 5, Appli
					Sequence 12, Appli
					Sequence 13, Appli
					Sequence 3, Appli
					Sequence 11, Appli
					Sequence 4, Appli
					Sequence 4, Appli
					Sequence 12, Appli
					Sequence 12, Appli
					Sequence 46, Appli
					Sequence 48, Appli
					Sequence 68, Appli

12	1624	53.1	568	13	US-10-173-519-2	Sequence 2, Appli
13	1624	53.1	568	15	US-10-403-161-72	Sequence 72, Appli
14	1624	53.1	568	16	US-10-490-080-1	Sequence 1, Appli
15	1624	53.1	568	17	US-10-718-359-6	Sequence 6, Appli
16	1622	53.0	568	15	US-10-092-900A-44	Sequence 44, Appli
17	1622	53.0	568	15	US-10-403-161-70	Sequence 70, Appli
18	1621	53.0	568	16	US-10-128-558-149	Sequence 149, Appli
19	1608	52.6	568	9	US-09-729-094-2	Sequence 2, Appli
20	1608	52.6	568	14	US-10-435-631-2	Sequence 2, Appli
21	1582	51.7	568	14	US-10-167-994-6	Sequence 6, Appli
22	1559.5	51.0	539	16	US-10-343-903-25	Sequence 25, Appli
23	1536.5	50.2	572	16	US-10-490-080-23	Sequence 23, Appli
24	1536.5	50.2	572	17	US-10-718-359-10	Sequence 10, Appli
25	1513.5	49.4	572	17	US-10-718-359-4	Sequence 4, Appli
26	1408	46.0	522	15	US-10-092-900A-50	Sequence 50, Appli
27	1408	46.0	522	15	US-10-403-161-74	Sequence 74, Appli
28	1377.5	45.0	600	14	US-10-167-994-13	Sequence 13, Appli
29	1377.5	45.0	600	17	US-10-718-359-14	Sequence 14, Appli
30	1363	44.6	602	14	US-10-167-994-4	Sequence 4, Appli
31	1337	43.7	516	15	US-10-092-900A-52	Sequence 52, Appli
32	1337	43.7	516	15	US-10-403-161-76	Sequence 76, Appli
33	1325	43.3	595	14	US-10-034-934-118	Sequence 118, Appli
34	1325	43.3	595	14	US-10-167-994-5	Sequence 5, Appli
35	1325	43.3	595	15	US-10-380-727-9	Sequence 9, Appli
36	1259	41.2	627	14	US-10-368-687-8	Sequence 8, Appli
37	1257	41.1	627	14	US-10-368-687-2	Sequence 2, Appli
38	1256	41.1	627	14	US-10-167-994-7	Sequence 7, Appli
39	1255.5	41.1	626	14	US-10-368-687-20	Sequence 20, Appli
40	1254	41.0	627	14	US-10-368-687-10	Sequence 10, Appli
41	1253.5	41.0	626	14	US-10-368-687-14	Sequence 14, Appli
42	1252	40.9	627	14	US-10-368-687-4	Sequence 4, Appli
43	1250.5	40.9	626	14	US-10-368-687-22	Sequence 22, Appli
44	1248.5	40.8	626	14	US-10-368-687-16	Sequence 16, Appli
45	1184.5	38.7	520	15	US-10-108-260A-3862	Sequence 3862, Appli
46	1097	35.9	581	14	US-10-368-687-12	Sequence 12, Appli
47	1095	35.8	581	14	US-10-368-687-6	Sequence 6, Appli
48	1093.5	35.8	580	14	US-10-368-687-24	Sequence 24, Appli
49	1091.5	35.7	580	14	US-10-368-687-18	Sequence 18, Appli
50	1042.5	34.1	572	14	US-10-167-994-2	Sequence 2, Appli
51	1042.5	34.1	572	17	US-10-017-479-2	Sequence 2, Appli
52	1042.5	34.1	572	17	US-10-718-359-2	Sequence 2, Appli
53	1039	34.0	382	16	US-10-128-558-346	Sequence 346, Appli
54	1014.5	33.2	407	11	US-09-833-245-1947	Sequence 1947, Appli
55	997	32.6	377	11	US-09-833-245-1949	Sequence 1949, Appli
56	887.5	29.0	551	17	US-10-718-359-8	Sequence 8, Appli
57	867	28.4	562	14	US-10-167-994-10	Sequence 10, Appli
58	849.5	27.8	585	13	US-10-173-519-4	Sequence 4, Appli
59	782	25.6	548	17	US-10-928-992-160	Sequence 160, Appli
60	781	25.5	552	10	US-09-882-227-496	Sequence 496, Appli
61	777	25.4	553	16	US-10-264-237-2022	Sequence 2022, Appli
62	776	25.4	553	16	US-10-437-963-127443	Sequence 127443, Appli
63	773	25.3	510	9	US-10-627-476-516	Sequence 516, Appli
64	765	25.0	548	16	US-09-738-626-3754	Sequence 3754, Appli
65	719	23.5	531	16	US-10-437-963-127444	Sequence 127444, Appli
66	706.5	23.1	304	15	US-10-425-115-250656	Sequence 250656, Appli
67	647	21.2	177	13	US-10-264-237-2022	Sequence 2022, Appli
68	516.5	16.9	189	15	US-10-173-519-8	Sequence 8, Appli
69	504	16.5	376	15	US-10-276-774-2634	Sequence 2634, Appli
70	493	16.1	233	15	US-10-424-599-254988	Sequence 254988, Appli
71	492.5	16.0	132	13	US-10-351-334-176	Sequence 176, Appli
72	490	16.0	446	15	US-10-173-519-7	Sequence 7, Appli
73	485	15.9	214	10	US-10-369-493-21334	Sequence 21334, Appli
74	485	15.9	214	10	US-09-974-879-354	Sequence 354, Appli
75	485	15.9	214	10	US-09-305-736-354	Sequence 354, Appli
76	485	15.9	214	11	US-09-818-683-354	Sequence 354, Appli
77	485	15.9	214	11	US-09-818-683-354	Sequence 354, Appli
78	484	15.8	146	15	US-10-621-401-354	Sequence 354, Appli
79	475	15.5	335	15	US-10-296-115-1187	Sequence 1187, Appli
80	407.5	13.3	518	17	US-10-627-476-518	Sequence 518, Appli
81	392	12.8	226	17	US-10-470-048B-584	Sequence 584, Appli
82	388.5	12.7	236	16	US-10-296-115-980	Sequence 980, Appli
83	364.5	11.9	453	15	US-10-767-701-43427	Sequence 43427, Appli
84	363.5	11.9	234	15	US-10-369-493-9936	Sequence 9936, Appli
					US-10-335-977-5397	Sequence 5397, Appli

85 358.5 11.7 162 11 US-09-833-245-1948 Sequence 1948, Ap  
86 350.5 11.5 162 11 US-09-833-245-1945 Sequence 1945, Ap  
87 336 11.0 1006 17 US-10-741-849-7011 Sequence 7011, Ap  
88 335.5 11.0 499 14 US-10-238-075-500 Sequence 500, App  
89 327 10.7 259 16 US-10-739-930-9458 Sequence 9458, Ap  
90 311.5 10.2 300 16 US-10-767-701-46317 Sequence 46317, A  
91 299 9.8 457 15 US-10-369-493-4266 Sequence 4266, Ap  
92 289 9.5 78 9 US-09-864-761-34252 Sequence 34252, A  
93 289 9.5 78 14 US-10-029-386-28038 Sequence 28038, A  
94 289 9.5 174 16 US-10-425-115-335649 Sequence 335649, A  
95 285 9.3 163 15 US-10-424-599-184881 Sequence 184881, A  
96 258.5 8.5 135 15 US-10-276-774-1592 Sequence 1592, Ap  
97 256.5 8.4 487 15 US-10-369-493-749 Sequence 749, App  
98 240.5 7.9 164 15 US-10-424-599-205152 Sequence 205152, A  
99 235.5 7.7 487 15 US-10-369-493-23559 Sequence 23559, A  
100 208 6.8 94 9 US-09-864-761-43642 Sequence 43642, A

ALIGNMENTS

RESULT 1  
US-10-173-519-5  
; Sequence 5, Application US/10173519  
; Publication No. US2002019582A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; TITLE OF INVENTION: 69624, A Human Transporter Family Member  
; TITLE OF INVENTION: and Uses Therefor  
; FILE REFERENCE: MP101-098P1RNM  
; CURRENT APPLICATION NUMBER: US/10/173,519  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: 60/298,970  
; PRIOR FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 587  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-173-519-5

Query Match 100.0%; Score 3058; DB 13; Length 587;  
Best Local Similarity 100.0%; Pred. No. 1.4e-280;  
Matches 587; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MATCPALWAYRFLYLIVLCPLIFLLPLPIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60  
DB 1 MATCPALWAYRFLYLIVLCPLIFLLPLPIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60  
QY 61 LFPDIVLPLMGINDASEVCIEYFKDTNLLFVGGLMVAIAVEHWNHAKRIALQVLLIIGVR 120  
DB 61 LFPDIVLPLMGINDASEVCIEYFKDTNLLFVGGLMVAIAVEHWNHAKRIALQVLLIIGVR 120  
QY 121 PALLLGLFMLVTAFLSMWISNTATTAMVPIGHAVLEQLQSKKDVGGNNNPTFELQEE 180  
DB 121 PALLLGLFMLVTAFLSMWISNTATTAMVPIGHAVLEQLQSKKDVGGNNNPTFELQEE 180  
QY 181 CPQKEVTKLDNGQPVSAFSEPRTKQTEHHRFSGLSLCICYSASIGGIATLTGTTNVLV 240  
DB 181 CPQKEVTKLDNGQPVSAFSEPRTKQTEHHRFSGLSLCICYSASIGGIATLTGTTNVLV 240  
QY 241 LOGQVNSLFPQNGVNVNFASFVGFAPPTMIILLLAWLQVLFVGNFRKNFGFGGEE 300  
DB 241 LOGQVNSLFPQNGVNVNFASFVGFAPPTMIILLLAWLQVLFVGNFRKNFGFGGEE 300  
QY 301 ERKQAAFOVKTQYRLGLGPMFAEKTIVLVLLVLMFTREPFGFGGDTVFANEKGQ 360  
DB 301 ERKQAAFOVKTQYRLGLGPMFAEKTIVLVLLVLMFTREPFGFGGDTVFANEKGQ 360  
QY 361 SMASDGTVAIFISLWVFIPSKIPGLMQDPKPKGLKAPPAILTWKTVNDKMPNIVILL 420

RESULT 2  
US-10-167-994-12  
; Sequence 12, Application US/10167994  
; Publication No. US20030082647A1  
; GENERAL INFORMATION:  
; APPLICANT: Reenan, Robert A.  
; APPLICANT: Rogina, Blanka  
; APPLICANT: Helfand, Stephen L.  
; TITLE OF INVENTION: TRANSPORTER PROTEIN  
; FILE REFERENCE: 13407-013001  
; CURRENT APPLICATION NUMBER: US/10/167,994  
; CURRENT FILING DATE: 2002-06-12  
; PRIOR APPLICATION NUMBER: US 60/255,013  
; PRIOR FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: US 10/017,479  
; PRIOR FILING DATE: 2001-12-12  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 587  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-167-994-12

Query Match 100.0%; Score 3058; DB 14; Length 587;  
Best Local Similarity 100.0%; Pred. No. 1.4e-280;  
Matches 587; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MATCPALWAYRFLYLIVLCPLIFLLPLPIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60  
DB 1 MATCPALWAYRFLYLIVLCPLIFLLPLPIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60  
QY 61 LFPDIVLPLMGINDASEVCIEYFKDTNLLFVGGLMVAIAVEHWNHAKRIALQVLLIIGVR 120  
DB 61 LFPDIVLPLMGINDASEVCIEYFKDTNLLFVGGLMVAIAVEHWNHAKRIALQVLLIIGVR 120  
QY 121 PALLLGLFMLVTAFLSMWISNTATTAMVPIGHAVLEQLQSKKDVGGNNNPTFELQEE 180  
DB 121 PALLLGLFMLVTAFLSMWISNTATTAMVPIGHAVLEQLQSKKDVGGNNNPTFELQEE 180  
QY 181 CPQKEVTKLDNGQPVSAFSEPRTKQTEHHRFSGLSLCICYSASIGGIATLTGTTNVLV 240  
DB 181 CPQKEVTKLDNGQPVSAFSEPRTKQTEHHRFSGLSLCICYSASIGGIATLTGTTNVLV 240  
QY 241 LOGQVNSLFPQNGVNVNFASFVGFAPPTMIILLLAWLQVLFVGNFRKNFGFGGEE 300  
DB 241 LOGQVNSLFPQNGVNVNFASFVGFAPPTMIILLLAWLQVLFVGNFRKNFGFGGEE 300  
QY 301 ERKQAAFOVKTQYRLGLGPMFAEKTIVLVLLVLMFTREPFGFGGDTVFANEKGQ 360  
DB 301 ERKQAAFOVKTQYRLGLGPMFAEKTIVLVLLVLMFTREPFGFGGDTVFANEKGQ 360  
QY 361 SMASDGTVAIFISLWVFIPSKIPGLMQDPKPKGLKAPPAILTWKTVNDKMPNIVILL 420  
DB 361 SMASDGTVAIFISLWVFIPSKIPGLMQDPKPKGLKAPPAILTWKTVNDKMPNIVILL 420  
QY 421 GGGFALAKGSEQSGLSLWGLDGLTFLQHIIPPSATAVILCLLIAIFTECTSNVATTTFLP 480

Db 421 GGGFALAKGSEGLSEWLGDKLTPLQHIPPSATAVILCLLIAIPTFECTSNVATTTFLP 480  
Qy 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAGFLNI 540  
Db 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAGFLNI 540  
Qy 541 IGVLAITLSINSWSIPFKLDTFPPSWAHNSNTSQCLLNPSNSTVPGGL 587  
Db 541 IGVLAITLSINSWSIPFKLDTFPPSWAHNSNTSQCLLNPSNSTVPGGL 587

RESULT 3  
US-10-718-359-13  
; Sequence 13, Application US/10718359  
; Publication No. US20050095240A1  
; GENERAL INFORMATION:  
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.  
; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION  
; FILE REFERENCE: 275.0008 0101  
; CURRENT APPLICATION NUMBER: US/10718,359  
; PRIOR FILING DATE: 2003-11-20  
; PRIOR FILING DATE: 2002-11-22  
; PRIOR FILING DATE: 2002-11-22  
; PRIOR FILING DATE: 2003-04-01  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 13  
; LENGTH: 587  
; TYPE: PRT  
; ORGANISM: rat NaDC1  
US-10-718-359-13

Query Match 100.0%; Score 3058; DB 17; Length 587;  
Best Local Similarity 100.0%; Pred. No. 1.4e-280;  
Matches 587; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATCHPALWAYRFLYLVLCLPIFLPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTA 60  
Db 1 MATCHPALWAYRFLYLVLCLPIFLPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTA 60

Qy 61 LFPVILFPLMGIMDASEVCIIEYFKDNTILFVGLMVAIAVEHWNHKLALQVLLIIGVR 120  
Db 61 LFPVILFPLMGIMDASEVCIIEYFKDNTILFVGLMVAIAVEHWNHKLALQVLLIIGVR 120

Qy 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLOQSKKDVGGNNPTFELQEE 180  
Db 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLOQSKKDVGGNNPTFELQEE 180

Qy 181 CPQKEVTKLDNGQPVSAPEPRTKTQEHHRFSQGLSLCICYSASIGGIATLTGTPNLV 240  
Db 181 CPQKEVTKLDNGQPVSAPEPRTKTQEHHRFSQGLSLCICYSASIGGIATLTGTPNLV 240

Qy 241 LQGVNSLPQNGNVNPFASWFGFAPPTMIILLLAWMLQVLFLGVNFRKNFGF 300  
Db 241 LQGVNSLPQNGNVNPFASWFGFAPPTMIILLLAWMLQVLFLGVNFRKNFGF 300

Qy 301 ERKQAFQVITQYRLGPMSPFAEKTIVTLFVLLVLMFTREPGFGDGVFANEGQ 360  
Db 301 ERKQAFQVITQYRLGPMSPFAEKTIVTLFVLLVLMFTREPGFGDGVFANEGQ 360

Qy 361 SNASDGTVAIFSLVMFIIPSKIPGLMOPKPKGLKAPPAILTWKTVNDKMPNIVILL 420  
Db 361 SNASDGTVAIFSLVMFIIPSKIPGLMOPKPKGLKAPPAILTWKTVNDKMPNIVILL 420

Qy 421 GGGFALAKGSEGLSEWLGDKLTPLQHIPPSATAVILCLLIAIPTFECTSNVATTTFLP 480  
Db 421 GGGFALAKGSEGLSEWLGDKLTPLQHIPPSATAVILCLLIAIPTFECTSNVATTTFLP 480

Qy 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAGFLNI 540  
Db 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAGFLNI 540

Qy 541 IGVLAITLSINSWSIPFKLDTFPPSWAHNSNTSQCLLNPSNSTVPGGL 587  
Db 541 IGVLAITLSINSWSIPFKLDTFPPSWAHNSNTSQCLLNPSNSTVPGGL 587

RESULT 4  
US-10-167-994-3  
; Sequence 3, Application US/10167994  
; Publication No. US20030082647A1  
; GENERAL INFORMATION:  
; APPLICANT: Reenan, Robert A.  
; APPLICANT: Regina, Blanka  
; APPLICANT: Helfand, Stephen L.  
; TITLE OF INVENTION: TRANSPORTER PROTEIN  
; FILE REFERENCE: 13407-013001  
; CURRENT APPLICATION NUMBER: US/10/167,994  
; CURRENT FILING DATE: 2002-06-12  
; PRIOR APPLICATION NUMBER: US 60/255,013  
; PRIOR FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: US 10/017,479  
; PRIOR FILING DATE: 2001-12-12  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 592  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-167-994-3

Query Match 78.2%; Score 2392; DB 14; Length 592;  
Best Local Similarity 76.9%; Pred. No. 2e-217;  
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;

Qy 1 MATCHPALWAYRFLYLVLCLPIFLPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTA 60  
Db 1 MATCHPALWAYRFLYLVLCLPIFLPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTA 60

Qy 61 LFPVILFPLMGIMDASEVCIIEYFKDNTILFVGLMVAIAVEHWNHKLALQVLLIIGVR 120  
Db 61 LFPVILFPLMGIMDASEVCIIEYFKDNTILFVGLMVAIAVEHWNHKLALQVLLIIGVR 120

Qy 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLOQSKKDVGGNNPTFELQ 178  
Db 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLOQSKKDVGGNNPTFELQ 178

Qy 179 EECPOKEVTKLDNGQ - PV-SAPSEPRTKTQEHHRFSQGLSLCICYSASIGGIATLTGT 235  
Db 179 EECPOKEVTKLDNGQ - PV-SAPSEPRTKTQEHHRFSQGLSLCICYSASIGGIATLTGT 235

Qy 236 TPNLVLOQVNSLPQNGNVNPFASWFGFAPPTMIILLLAWMLQVLFLGVNFRKNFGF 295  
Db 236 TPNLVLOQVNSLPQNGNVNPFASWFGFAPPTMIILLLAWMLQVLFLGVNFRKNFGF 295

Qy 296 GEGEERKQAFQVITQYRLGPMSPFAEKTIVTLFVLLVLMFTREPGFGDGVFANEGQ 355  
Db 296 GEGEERKQAFQVITQYRLGPMSPFAEKTIVTLFVLLVLMFTREPGFGDGVFANEGQ 355

Qy 356 NEKQSMASDGTVAIFSLVMFIIPSKIPGLMOPKPKGLKAPPAILTWKTVNDKMPN 415  
Db 356 NEKQSMASDGTVAIFSLVMFIIPSKIPGLMOPKPKGLKAPPAILTWKTVNDKMPN 415

Qy 416 IVILLGGGFALAKGSEGLSEWLGDKLTPLQHIPPSATAVILCLLIAIPTFECTSNVAT 475  
Db 416 IVILLGGGFALAKGSEGLSEWLGDKLTPLQHIPPSATAVILCLLIAIPTFECTSNVAT 475

Qy 476 TFLPILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAG 535  
Db 476 TFLPILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAG 535

Qy 536 FLNIIIGVLAITLSINSWSIPFKLDTFPPSWAHNSNTSQCLLNPSNSTVPGGL 584  
Db 536 FLNIIIGVLAITLSINSWSIPFKLDTFPPSWAHNSNTSQCLLNPSNSTVPGGL 584

```

RESULT 5
US-10-167-994-11
; Sequence 11, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-167-994-11

Query Match      78.2%; Score 2392; DB 14; Length 592;
Best Local Similarity 76.9%; Pred. No. 2e-217;
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;

QY 1 MATCPALMAYRFLIIVLCPIFLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60
DB 1 MATCQALMAYRSYLIVFFPILLPLILVPSKEAYCAVAIILMALFWCTEALPLAVTA 60

QY 61 LPFIVLPALMIMDASEVCIEYFKDNTILFVGGMLVAIAVEHNLHRIALQVLLIGVR 120
DB 61 LPFILFPMGIVDASEVAYEYLKSNLLFPFGLLVAIAVEHNLHRIALRVLLIGVR 120

QY 121 PALILIGFMLVATFLSWISNTATAMVPIGHAVLEOLQSK--KDVGGNNPTFELQ 178
DB 121 PAPILIGFMLVATFLSWISNTATAMVPIGHAVLEOLQSK--KDVGGNNPTFELQ 180

QY 179 EECQKEVTKLDNGQ--PV-SAPSEPRTKTQEHHRFSQGLSLCICYSASIGGIATLTGT 235
DB 181 EPSQKEVTKLDNGQALPVTISASSEGRAHLSQKHLTLTQCMSLCVCYSASIGGIATLTGT 240

QY 236 TPNLVLOQVNSLFPQNGNVNFAFWGFPAPPTMIILLIALLWLVLFVGNFRKNFGF 295
DB 241 APNLVLOQVNSLFPQNGNVNFAFWGFPAPPTMIILLIALLWLVLFVGNFRKNFGI 300

QY 296 GEGEERKQAAFOVIKTOYRLLGPMSPAEKTVTVLVLLVLMFTREPGFPFGMGDTVFA 355
DB 301 GEKQEQOQAAVCVQTEHRLGLGPMTPAEKATISILFVILVLLWTRBPGFPLGNGLAFP 360

QY 356 NEKQSMASDGTVAIFISLVNFIIPSKIPLGMQDPKPGKLKAPPAILTWTNDKMPWN 415
DB 361 NAKGESMVSDGTVAIFIGIIMFIIPSKPPGLTQDPENPGKLKAPGLDLDKWTNQKMPWN 420

QY 416 IVILLGGFALAKSQSGSLSEWLDKLTPTLOHTPPSATAVILCLLAIETECTSNVATT 475
DB 421 IVLLGGGYALAKGSESGSLSEWLGKLTPTLOSPAPAPAIAIILSLLVATFTECTSNVATT 480

QY 476 TLEPLILASMAQAICLHPLYVNLPCITLAASLAFMLPVATPPNAIVFSGGLKVDMDARAG 535
DB 481 TIFLPLILASMAQAICLHPLYVNLPCITLATSALFMLPVATPPNAIVFSEGLKVLDMARAG 540

QY 536 FLLNIIGVLAITLSINSWSIPKLDTPPSWAHNS-TSQCLLNPSNSTVP 584
DB 541 FLLNIIGVLIILAINSGIPLFSLHSPSPSAQSNNTTAQCLPSLANITTP 590

RESULT 6
US-09-729-094-4
; Sequence 4, Application US/09729094
; GENERAL INFORMATION:
; APPLICANT: CHATURVEDI, Kabir et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: CL000662
; CURRENT APPLICATION NUMBER: US/09/729,094
; CURRENT FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-729-094-4

Query Match      64.5%; Score 1971.5; DB 9; Length 619;
Best Local Similarity 63.2%; Pred. No. 1.6e-177;
Matches 386; Conservative 87; Mismatches 85; Indels 53; Gaps 10;

QY 10 AYRFYLVLCPLIFLPLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTALFVLPFL 69
DB 10 ANRNYFIPLVPLFLPLPLVVPVPTKEASCGFVIVNALFWCTEALPLAVTALFVLPFL 69

QY 70 MGIMDASEVCIEYFKDNTILFVGGMLVAIAVEHNLHRIALQVLLIGVRPALLLGM 129
DB 70 MGIMDSTAVCSQYLKDTNMLFVGGMLVAIAVEHNLHRIALRVLLIGVVKPALLLGM 129

QY 130 LVTAFLSMWISNTATAMVPIGHAVLEOLQSK-----KDVGGNN----- 171
DB 130 VVTAFLSMWISNTATAMVPIGHAVLEOLQSK-----KDVGGNN----- 189

QY 172 -----NPTFELQEEC-----POKEVTKLDNGQPVSAAP-SEPRTOKT 206
DB 190 SVNPSGKMLAINDTATVATENEGFEIQEKTKDPEKQ-EKQSIGIVIEPEDEKQTEK 248

QY 207 Q--EHRFSQGLSLCICYSASIGGIATLTGTPNLVLOQVNSLFPQNGNVNFAFWGFG 264
DB 249 QKEKHLKICKGMSLVCYSASIGGIATLTGTPNLVLOQVNSLFPQNGNVNFAFWGFG 308

QY 265 APPTMIILLIALLWLVLFVGNFRKNFGEGEER--ERKQAAFOVIKTOYRLLGPMSP 322
DB 309 APPTMLVLLALSMLWLFQTYLVGNFKKNGFCGCGNAEQEKEKRAFRVISEGHKKLGSMTF 368

QY 323 AEKTVTVLVLLVLMFTREPGFPFGMGDTVFAVNEKQSMASDGTVAIFISLVNFIIPSK 382
DB 369 AEISVLVLFVLLVLMFTREPGFPFGMA-TISFNKGGKEWMTDATVAIFVSLAMFFPPE 427

QY 383 IPGL-MQDPKPG---KLKAPPAILTWTNDKMPWNIVILLGGGFALAKSQSGSLSEW 438
DB 428 LPSFKYQDTPKPGKPKLRVPPALLDWKTVNEKMPWNIVILLGGGFALAKSQSGSLSLW 487

QY 439 LGDKLTPLOHIPPSTAVTILCLLAIETECTSNVATTTLFPLILASMAQAICLHPLYVNL 498
DB 488 LGSEKLPLOSIPPAIALILCLLVATFTECTSNVATTTLFPLILASMAQAICLHPLYVNL 547

QY 499 PCTLAASLAFMLPVATPPNAIVFSGGLKVDMDARAGFLNIITGVLAITLSINSWSIPF 558
DB 548 PCTLSLAFMLPVATPPNAIVFSGGLKVDMDARAGFLNIITGVLAITLSINSWSIPF 607

QY 559 KLDTPPSWAHS 569
DB 608 NLGTFPSWANA 618

RESULT 7
US-10-435-631-4
; Sequence 4, Application US/10435631
; Publication No. US20030186381A1
; GENERAL INFORMATION:
; APPLICANT: CHATURVEDI, Kabir et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,

```

```

; ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; ; TITLE OF INVENTION: AND USES THEREOF
; ; FILE REFERENCE: CL0006G2CON
; ; CURRENT APPLICATION NUMBER: US/10/435,631
; ; CURRENT FILING DATE: 2003-05-12
; ; NUMBER OF SEQ ID NOS: 4
; ; SOFTWARE: FastSeq for Windows Version 4.0
; ; SEQ ID NO 4
; ; LENGTH: 619
; ; TYPE: PRT
; ; ORGANISM: Xenopus laevis
US-10-435-631-4

```

Query Match 64.5%; Score 1971.5; DB 14; Length 619;  
Best Local Similarity 63.2%; Pred. No. 1.6e-177;  
Matches 386; Conservative 87; Mismatches 85; Indels 53;  
Gaps 10;

10	QY	AVRFYILVLCPIFLPLPLVQTKQYKAYSIIMALLMCTEALPLAVTALFPVILFPL	69
10	Db	ANRNYFIIFLPLFLPLPLVPTKEASCQGVIIIMALFWCTEALPLAVTALFPVLLPDM	69
70	QY	MGIMDASEVCLEYKQTNILPFGVLMAVAIEVHNHLKRIALQVILIIIVRPALLILGPM	129
70	Db	MGIMDSTAVCQOYLKQTNMLFISGILLVAISVEKWNHLKRIALRVLLIVGVKPALLLGPM	129
130	QY	LVTAPLSMWISNTATTAMVPVIGHAVLEQLQGSK---KDVEGNN-----	171
130	Db	VVTAPLSMWISNTATTAMVPIAQAVMEQLHSSEKVDERVEGNSNTQKNVNCMDMYE	189
172	QY	-----NPTFELQSEC-----POKEVTKLNGQPVISAP-SEPTQKT	206
190	Db	SYMPSGKNALAIQNTYATENEGFIEQKSTKDPESKQ-EKQSIGIVIEPEDEKQTEK	248
207	QY	Q---EHHREFSGLSLICVCYSASIGGIATLTGTTNPLVLQGVNSLFPONGNVNVEASWCF	264
249	Db	QKERHLKICKNSLVCVCYSASIGGIATLTGTTNPLVNGQMDSELPENNNIINFASWFG	308
265	QY	AFPTMIILLLLAWLQVLFLGVNFRKNFGFGESE--ERKQAQFOVIKTYRLLAGPMSF	322
309	Db	AFPTMLVLLALSLSWLQVILYGVNFKNFCGGNAEQEKEKEKAPRVISGEHKLGSMTF	368
323	QY	AEKTVTVLVLVLLVNLWTRBPGPGHGDIVFANEKGQSWASDGTVAIFISLWFIIPSK	382
369	Db	AEISVLVLIILLVLLWTRBPGPGHWA-TISFNKGGKEMVTDATVAIFVSLMFFPFE	427
383	QY	IPGL-WQDPKPKPG---KLKAPPAILTKWTVNDKMPNIVILGGGFALAKGSEQGLSEW	438
428	Db	LPSFKYQDTRPGMKPKLRVPVALLDWKTNKKMPNIVILGGGFALAKGSEEGSLW	487
439	QY	LGDKLTPLOHTPPSATVILCLLIAIETECTSNVATTTFLPLPLASMAQAICLHPLYVWL	498
488	Db	LGEKLPILQSIPPAAIALIICLVATFECTSNVATTTFLPLPLASMAKAIQLNPLYIML	547
499	QY	PCTLAASLAFMLPVATPPNAIVSFSGLVKYSMDARAGFLNIIGVLAITLSNNSIPF	558
548	Db	PCTLSASLAFMLPVATPPNAIAFSYGOQKVIDMAKAGLLNIIIGVLTITLANSMGTFM	607
559	QY	KLDTFPFSWAHS	569
608	Db	NIGTFPFSWANA	618

RESULT 8  
US-10-718-359-12  
; Sequence 12, Application US/10718359  
; Publication No. US20050095240A1  
; GENERAL INFORMATION:  
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.  
; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION  
; FILE REFERENCE: 275, 0008 0101  
; CURRENT APPLICATION NUMBER: US/10/718,359  
; CURRENT FILING DATE: 2003-11-20  
; PRIOR APPLICATION NUMBER: 60/428,469

```

; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 60/459,441
; PRIOR FILING DATE: 2003-04-01
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 561
; TYPE: PRT
; ORGANISM: zebrafish NaCT
US-10-718-359-12

```

Query Match	59.8%	Score 1829;	DB 17;	Length 581;
Best Local Similarity	60.4%;	Pred. No. 4.6e-164;		
Matches 349;	Conservative 92;	Mismatches 113;	Indels 24;	Gaps 6
QY	8	LWAYRFVILVLCPIELLPLPLIVOTKEAYCAYSIILMALLWCTEALPLAVTALFPVILF	67	
DB	9	VWVKWTLILFCTPFIPLPLPLVIGSKEACCAVVVLMAVTCVTEVPLAVTALLPAVLF	68	
QY	68	PLMGIMDASEVCIEYFKDNIPLFVGLGMAVAIVEHWNHLKRIALQVLLIIGVRPALLLG	127	
DB	69	PLPRIMESQDVNCQYLKDTNMLPLGLGLMAVAIVEHWNHLKRIALRVLLLVGVRPALLMLG	128	
QY	128	FMILVTAFLSNWISNTATTAMWVPIGHAVLEOLOGSKDVEGG---NNNPTELQEBCEQK	184	
DB	129	FMGVTAFLSNWISNTATTAMWVPIQVAVLEQLNNTAQOBOSISPETEKEKQEPESGPE	188	
QY	185	EVTKLNGQVPVASPSEP--RTQKTQBHHFRSQGLSCLICYSASIGGTATLTGTPNLVLQ	242	
DB	189	EKVVL-NGDNFMSQDEEHSREAEERLKMSEGLTLCVCVAASIGGTATLTGTCPNVLVM	247	
QY	243	GQVNSLFPQNGNVNVPASWFGAFPPTMIILLLAWLQVLFLGVNPRKNPFGQGBEER	302	
DB	248	GQMSQLFPDNPDIINPASFVGFAPFNMIIMLTAWLWLQVFLGINFKKTGWCGCTVATEK	307	
QY	303	KQAAFOVITQVRLGLGPMSPAEKTVTVLVLVVLVMTRETPGPPGWDVTPANEKQSQM	362	
DB	308	ETAAVNVIIKEEHSGLGPMTFGELUSVLALFLLVVLWPTRPDGFVDGWATRFNNADK--EF	365	
QY	363	ASDGTVAIFISLVMFIPIKIPQLM-----QDPKKPKKLKAPPAILTWKTVNDK	411	
DB	366	VTDAVAVFVAALLFVPSKPPRLCFMRTESEFDTVPQESGP-----TPALLTWKVTKK	420	
QY	412	MPWNVILLGGGFALAKGSEQGLSWLGDGLTPLQHIPPSATAVILCLIIAIPTECTSN	471	
DB	421	MPWSIILLGGGFALAKGSEISGLSKWLGDQMSPLQSIIPWATAIVICLMIATPTECTSN	480	
QY	472	VATTTFLPLILASMAQAI CLHPLVYMLPCTTLAASLAPMLPVATPPNAIVFSPGLKVKYSDM	531	
DB	481	VATATLFLPLASMSQSIGVNPVYVWVPTLSASFAPMLPVATPPNAIVFSGYLKVKYSDM	540	
QY	532	ARAGFLNIIIGVLATILTSINSWSIPIFKLDTPPSWAHS	569	
DB	541	AKTGIVMNIIGLISITLAINSGRAIFSLDTPPSWANT	578	

RESULT 9

US-10-092-900A-46

Sequence 46, Application US/10092900A

Publication No. US20040043382A1

GENERAL INFORMATION:

APPLICANT: Padigaru, Muralidhara

APPLICANT: Spytek, Kimberly A.

APPLICANT: Shenoy, Suresh G.

APPLICANT: Taupier Jr., Raymond J.

APPLICANT: Pena, Carol E.A.

APPLICANT: Li, Li

APPLICANT: Zerhuseen, Bryan D.

APPLICANT: Gusev, Vladimir Y.

APPLICANT: Ji, Weizhen

APPLICANT: Gorman, Linda

APPLICANT: Miller, Charles E.

APPLICANT: Kekuda, Ramesh

Db	241	QCQDERKRLCKAMTICICVAAASIGGTATLTGTGPNVLLGQNNELFPDSDKLVNFPASW	300
QY	262	FGFAPFTMIILLLLAWLQWLFLGVNPRKNFGGEEBEERKQAAQVIKTOYRLLGPM	321
Db	301	FAFAFENMLVLLFAWLQFVYMFSSFKSGGCGLESKKNEKAALKVLQBEYKRLGPLS	360
QY	322	FAEKTVTVLVLLVLTFTBPGPPGPGWDTVFANEKGQSMASDGTVAIFISLVMFIP	381
Db	361	FAEINVLICFFLLVLTNFSRDPGFNPGWLTVAWVEGETKSV-SDATVAIFATLLFIVPS	419
QY	382	KIP-----GLMDPKKPKGLKAPPAILTWKTVNDKMPNIVILGGGFPALAKGSEQSGLS	436
Db	420	QXPKFNERSQTEEGKSP-VLIAPPPLLDWKVQEPKPGIIVLLGGGFPALAKGSEASGLS	478
QY	437	EWLQKLTPLQHIHPSATAVILCLIAIETCTSNVATTLFLPLTLASMAQAICLHPLYV	496
Db	479	VVMGQMEPLHAPVPAATLILSLVAVFETCTSNVATTLFLPLTFASMSKSIQUNPLYI	538
QY	497	MLPCTLASLAFMLPVATPPNAIVFSFGGLKVDNARAGFLNIIGVLAITLSINSNIP	556
Db	539	MLPCTLSASFAPMLPVATPPNAIVFTYGHLLKVADWVKTVIMNIIIGVFCVFLAVNTWGRA	598
QY	557	IFKLDTFPSWAH 568	
Db	599	IFDLDFPDWAN 610	

RESULT 10

US-10-092-900A-48

Sequence 48, Application US/10092900A

Publication No. US20040043382A1

GENERAL INFORMATION:

APPLICANT: Padigar, Muralidhara

APPLICANT: Spytek, Kimberly A.

APPLICANT: Shenoy, Suresh G.

APPLICANT: Taupier Jr., Raymond J.

APPLICANT: Pena, Carol E.A.

APPLICANT: Li, Li

APPLICANT: Zerhusen, Bryan D.

APPLICANT: Gusev, Vladimir Y.

APPLICANT: Ji, Weizhen

APPLICANT: Gorman, Linda

APPLICANT: Miller, Charles E.

APPLICANT: Kekuda, Ramesh

APPLICANT: Patturajan, Meera

APPLICANT: Gangolli, Esha A.

APPLICANT: Vernet, Corine A.M.

APPLICANT: Guo, Xiaojia Sasha

APPLICANT: Tchernev,, Velizar T.

APPLICANT: Fernandes, Elma R.

APPLICANT: Casman, Stacie J.

APPLICANT: Malyankar, Uriel M.

APPLICANT: Gerlach, Valerie

APPLICANT: Liu, Yi

APPLICANT: Anderson, David W.

APPLICANT: Spaderna, Steven K.

APPLICANT: Catterton, Elina

APPLICANT: Leite, Mario W.

APPLICANT: Zhong, Haihong

APPLICANT: Alsobrook, John P.

APPLICANT: Lepley, Denise M.

APPLICANT: Rieger, Daniel K.

APPLICANT: Burgess, Catherine E.

TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same

FILE REFERENCE: 21402-290C

CURRENT APPLICATION NUMBER: US/10/092,900A

CURRENT FILING DATE: 2002-03-07

PRIOR APPLICATION NUMBER: USSN 60/274,322

PRIOR FILING DATE: 2001-03-08

PRIOR APPLICATION NUMBER: USSN 60/283,675

PRIOR FILING DATE: 2001-04-13

PRIOR APPLICATION NUMBER: USSN 60/338,092

PRIOR FILING DATE: 2001-12-03

```
/ PRIOR APPLICATION NUMBER: USSN 60/274,281
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: USSN 60/274,191
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: USSN 60/325,681
/ PRIOR FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: USSN 60/304,354
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: USSN 60/279,995
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: USSN 60/294,899
/ PRIOR FILING DATE: 2001-05-31
/ PRIOR APPLICATION NUMBER: USSN 60/287,424
/ PRIOR FILING DATE: 2001-04-30
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 768
/ SEQ ID NO 48
/ LENGTH: 616
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-092-900A-48

Query Match
Best Local Similarity 53.2%; Score 1628; DB 15; Length 616;
Matches 314; Conservative 113; Mismatches 139; Indels 46; Gaps 7;

Qy 1 MATCHPALWAYRYLVLCPLIFLLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTA 60
Db 1 MASALSYVSKPKSVFLFVTPLELLPLVILMPAKVSCCAYVILMAIYVCTEVIPLAVTS 60
Qy 61 LPPIVLPMLGMDASEVCIEYFKDTNLLFVGLGMVAIAVEHNLHKKRIALQVLLIIGVR 120
Db 61 LMPVLLPFLQILDSRQVCQVMKDTNMLFLGGLIVAVAVERNLHKKRIALRTLLWVGAK 120
Qy 121 PALLLGLMVLTAFLSMWISNTATTAMVPIGHAVLEQLQSGKDVGG-----NNN 172
Db 121 PARMLGFMGVLTALLSMWISNTATTAMVPIVEAILQMEATSAATEAGLEGQGTINNL 180
Qy 173 PTFE-----LOEBCPQ-----KEVTKL-----DNGQPVSAPEP 201
Db 181 NALEDDETVAIVLGKCVAILSTYVKKVEKLQINMLTPKLEKQEQDGLGPIRQDSA 240
Qy 202 RTQKTOEHHRFSQGLSLCICYSASIGGIATLTGTTNPLVLOGVNSLFPQNGNVNPFASW 261
Db 241 QCQEDQERKRLCKAMTLCICVAASIGGTATLTGTGPNVLLQGMNELFPDSKDLVNFASW 300
Qy 262 FGPAFTMIILLALLLAWLQVLFVGNFRKNFGFGEGERKQAAFOVIKTOYRLLGPM 321
Db 301 FAFAPNNMLVMLLFAWLQVLYMFSSFKSGWGLSKKNEKAAALKVLOEYRKLGLPLS 360
Qy 382 KIP-----GLMQDPKKPKGLKAPPAILTWTNDKMPNIVILLGGGFALAKGSEQSGLS 436
Db 420 QPKFENFRSQTGEGKSP-VLIAPPPLLDWKVQKVPNGVILLGGGFALAKGSEASGLS 478
Qy 437 EWLGDKLTPLOHIPSATAVILCLLIAIPTCTSNVATTTFLPLTASMAOAILCHPLIYV 496
Db 479 VMGQKMEPLHAPVPAATLTLSSLVAVFTCTSNVATTTFLPLIFASMSRSIGLNPLYI 538
Qy 497 MLPCTLAASLAFMLPVATPPNAIVPSFGGLKVSMDARAGFLNIIIGVLAITLSINSWSIP 556
Db 539 MLPCTLSASFAMLPVATPPNAIVFTYGHKLKADVMKTVGIMNIIIGVFCVFLAVNTWGRA 598
Qy 557 IFKLDTPFSWAH 568
Db 599 IFDLDPFDWAN 610
```

RESULT 11  
US-10-403-161-68

```
/ Sequence 68, Application US/10403161
/ Publication No. US20040043930A1
/ GENERAL INFORMATION:
/ APPLICANT: Anderson, David et al.
/ TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 21402-573C
/ CURRENT APPLICATION NUMBER: US/10/403,161
/ CURRENT FILING DATE: 2003-03-31
/ PRIOR APPLICATION NUMBER: 60/370349
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: 60/384543
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: 60/370969
/ PRIOR FILING DATE: 2002-04-08
/ PRIOR APPLICATION NUMBER: 60/403748
/ PRIOR FILING DATE: 2002-08-15
/ PRIOR APPLICATION NUMBER: 60/372019
/ PRIOR FILING DATE: 2002-04-12
/ PRIOR APPLICATION NUMBER: 60/374379
/ PRIOR FILING DATE: 2002-04-22
/ PRIOR APPLICATION NUMBER: 09/779679
/ PRIOR FILING DATE: 2001-02-08
/ PRIOR APPLICATION NUMBER: 60/181045
/ PRIOR FILING DATE: 2000-02-08
/ PRIOR APPLICATION NUMBER: 10/055877
/ PRIOR FILING DATE: 2002-01-22
/ PRIOR APPLICATION NUMBER: 60/262892
/ PRIOR FILING DATE: 2001-01-19
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 173
/ SOFTWARE: Curaseqlist version 0.1
/ SEQ ID NO 68
/ LENGTH: 616
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-403-161-68

Query Match
Best Local Similarity 53.2%; Score 1628; DB 15; Length 616;
Matches 314; Conservative 113; Mismatches 139; Indels 46; Gaps 7;

Qy 1 MATCHPALWAYRYLVLCPLIFLLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTA 60
Db 1 MASALSYVSKPKSVFLFVTPLELLPLVILMPAKVSCCAYVILMAIYVCTEVIPLAVTS 60
Qy 61 LPPIVLPMLGMDASEVCIEYFKDTNLLFVGLGMVAIAVEHNLHKKRIALQVLLIIGVR 120
Db 61 LMPVLLPFLQILDSRQVCQVMKDTNMLFLGGLIVAVAVERNLHKKRIALRTLLWVGAK 120
Qy 121 PALLLGLMVLTAFLSMWISNTATTAMVPIGHAVLEQLQSGKDVGG-----NNN 172
Db 121 PARMLGFMGVLTALLSMWISNTATTAMVPIVEAILQMEATSAATEAGLEGQGTINNL 180
Qy 173 PTFE-----LOEBCPQ-----KEVTKL-----DNGQPVSAPEP 201
Db 181 NALEDDETVAIVLGKCVAILSTYVKKVEKLQINMLTPKLEKQEQDGLGPIRQDSA 240
Qy 202 RTQKTOEHHRFSQGLSLCICYSASIGGIATLTGTTNPLVLOGVNSLFPQNGNVNPFASW 261
Db 241 QCQEDQERKRLCKAMTLCICVAASIGGTATLTGTGPNVLLQGMNELFPDSKDLVNFASW 300
Qy 262 FGPAFTMIILLALLLAWLQVLFVGNFRKNFGFGEGERKQAAFOVIKTOYRLLGPM 321
Db 301 FAFAPNNMLVMLLFAWLQVLYMFSSFKSGWGLSKKNEKAAALKVLOEYRKLGLPLS 360
Qy 382 KIP-----GLMQDPKKPKGLKAPPAILTWTNDKMPNIVILLGGGFALAKGSEQSGLS 436
Db 420 QPKFENFRSQTGEGKSP-VLIAPPPLLDWKVQKVPNGVILLGGGFALAKGSEASGLS 478
```

QY 437 EWLGDKLTPLOHIPPATAVILCLLIAITFECTSNVATTTFLPILASMAQAICLHPLV 496  
 Db 479 VMGKQMEPLHAPPAITLILSVAVTECHSNVATTTFLPILASMSRSICLNPLV 538  
 QY 497 MLPCTLLAASLAFMLPVATPPNAIVFSGGLKVSDMARAGFLNIIIGVLAITLINSWSIP 556  
 Db 539 MLPCTLSASFAFMLPVATPPNAIVFTYGHKLVADVMKTVGIMNIIIGVCFVLAVENTWGRA 598  
 QY 557 IFKLDTFPPSWAH 568  
 Db 599 IFDLDFPDWAN 610

RESULT 12  
 US-10-173-519-2  
 ; Sequence 2, Application US/10173519  
 ; Publication No. US20020193582A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Curtis, Rory A.J.  
 ; TITLE OF INVENTION: 59624, A Human Transporter Family Member  
 ; TITLE OF INVENTION: and Uses Therefor  
 ; FILE REFERENCE: MPI01-098PINNM  
 ; CURRENT FILING DATE: 2002-06-17  
 ; PRIOR APPLICATION NUMBER: 60/298, 970  
 ; PRIOR FILING DATE: 2001-06-18  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 568  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-173-519-2

Query Match 53.1%; Score 1624; DB 13; Length 568;  
 Best Local Similarity 53.2%; Pred. No. 1.2e-144;  
 Matches 309; Conservative 108; Mismatches 132; Indels 32; Gaps 6;

QY 1 MATCWPAALWAYRFLIVLCIPFLPLPLIVOTKEAYCAYSIILMALLWCTEALPLAVTA 60  
 Db 1 MASALSYSVKFSFVLFTVPLLLPLVILMPAKFVRCAYVILMAIYKTEVIPLAVTS 60  
 QY 61 LPFIVLPFLMGINDASEVCIEYKOTNIFVGGIMVAIAVEHNLKRIALQVLLIIGVR 120  
 Db 61 LMPVLLPFLQILDSRQVCQVMKTNMLFLGLLIIVAVAVERNLKHRIALRTLLWVGAK 120  
 QY 121 PALLLGFMVLVTAFLSMWISNTATTAMVPPIGHAVLEQLOGSKKDVGGNNPTFELQEE 180  
 Db 121 PARMLGFMGVMTALLSMWISNTATTAMVPPIVEAILQOMEATSAATEAG-----LELVDK 175  
 QY 181 CPQKEVTKLNDGQVPSAPSEPRKTQEHHRFSGGLSLCICYSASIGGIATLTGTPNLV 240  
 Db 176 GKAKE---LPGSQVIFEGPTLQGEDQERKLCCKAMTLCICYAASIGGTATLTGTGNV 232  
 QY 241 LQGVNSLFPQNGNVNFAFWGFAFPTMILLLLWMLQVLFVGNFRKNFGFGGEE 300  
 Db 233 LLGQNNELFPDSKDLNFAFWGFAFPMNLVMLLFAWLMLQFYVMRNFKKSGCGLESK 292  
 QY 301 ERKQAAFOVTKTVRLGPMSPFAEKTVTVLVLLVWLFTRPFGPGW-----GDTV 353  
 Db 293 KNEKAALKVQEEYRKLGPLSFASINVLICFFLLVILWFSRDPGFMPLVAVVEGETK 352  
 QY 354 FANEKQGSMDGTVAFISVLVMIIPKIPGL-----MDPKPKGLKAPPAITLTKT 407  
 Db 353 Y-----VSDATVAIFVATLLFVPSQPKFNFRSQTEERTKTP---FYPPPLLDKW 401  
 QY 408 VNDKMPNIVILLGGGALAKGSQSGLEWLGDKLTPLOHIPPATAVILCLLIAIFTE 467  
 Db 402 TOEKVPWGVILLGGGALAKGSASGLSVWVGQMEPLHAPPAITLILSVAVFTE 461  
 QY 468 CTSNVATTTFLPILASMAQAICLHPLVYMLPCTLLAASLAFMLPVATPPNAIVFSGGLK 527

Db 462 CTSNVATTTFLPILASMSRSICLNPLVIMLPCTLSASFAFMLPVATPPNAIVFTYGHK 521  
 QY 528 VSDMARAGFLNIIIGVLAITLINSWSIPIFKLDTFPPSWAH 568  
 Db 522 VADMVKTGVMNIIIGVCFVLAVENTWGRAIFOLDHFPDWA 562

RESULT 13  
 US-10-403-161-72  
 ; Sequence 72, Application US/10403161  
 ; Publication No. US2004004390A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Anderson, David et al.  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: 21402-573C  
 ; CURRENT APPLICATION NUMBER: US/10/403,161  
 ; CURRENT FILING DATE: 2003-03-31  
 ; PRIOR APPLICATION NUMBER: 60/370349  
 ; PRIOR FILING DATE: 2002-04-05  
 ; PRIOR APPLICATION NUMBER: 60/384543  
 ; PRIOR FILING DATE: 2002-05-30  
 ; PRIOR APPLICATION NUMBER: 60/370969  
 ; PRIOR FILING DATE: 2002-04-08  
 ; PRIOR APPLICATION NUMBER: 60/403748  
 ; PRIOR FILING DATE: 2002-08-15  
 ; PRIOR APPLICATION NUMBER: 60/372019  
 ; PRIOR FILING DATE: 2002-04-12  
 ; PRIOR APPLICATION NUMBER: 60/374379  
 ; PRIOR FILING DATE: 2002-04-22  
 ; PRIOR APPLICATION NUMBER: 09/779679  
 ; PRIOR FILING DATE: 2001-02-08  
 ; PRIOR APPLICATION NUMBER: 60/181045  
 ; PRIOR FILING DATE: 2000-02-08  
 ; PRIOR APPLICATION NUMBER: 10/055877  
 ; PRIOR FILING DATE: 2002-01-22  
 ; PRIOR APPLICATION NUMBER: 60/262892  
 ; PRIOR FILING DATE: 2001-01-19  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 173  
 ; SOFTWARE: CuraSeqList version 0.1  
 ; SEQ ID NO 72  
 ; LENGTH: 568  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-403-161-72

Query Match 53.1%; Score 1624; DB 15; Length 568;  
 Best Local Similarity 53.2%; Pred. No. 1.2e-144;  
 Matches 309; Conservative 108; Mismatches 132; Indels 32; Gaps 6;

QY 1 MATCWPAALWAYRFLIVLCIPFLPLPLIVOTKEAYCAYSIILMALLWCTEALPLAVTA 60  
 Db 1 MASALSYSVKFSFVLFTVPLLLPLVILMPAKFVRCAYVILMAIYKTEVIPLAVTS 60  
 QY 61 LPFIVLPFLMGINDASEVCIEYKOTNIFVGGIMVAIAVEHNLKRIALQVLLIIGVR 120  
 Db 61 LMPVLLPFLQILDSRQVCQVMKTNMLFLGLLIIVAVAVERNLKHRIALRTLLWVGAK 120  
 QY 121 PALLLGFMVLVTAFLSMWISNTATTAMVPPIGHAVLEQLOGSKKDVGGNNPTFELQEE 180  
 Db 121 PARMLGFMGVMTALLSMWISNTATTAMVPPIVEAILQOMEATSAATEAG-----LELVDK 175  
 QY 181 CPQKEVTKLNDGQVPSAPSEPRKTQEHHRFSGGLSLCICYSASIGGIATLTGTPNLV 240  
 Db 176 GKAKE---LPGSQVIFEGPTLQGEDQERKLCCKAMTLCICYAASIGGTATLTGTGNV 232  
 QY 241 LQGVNSLFPQNGNVNFAFWGFAFPTMILLLLWMLQVLFVGNFRKNFGFGGEE 300  
 Db 233 LLGQNNELFPDSKDLNFAFWGFAFPMNLVMLLFAWLMLQFYVMRNFKKSGCGLESK 292  
 QY 301 ERKQAAFOVTKTVRLGPMSPFAEKTVTVLVLLVWLFTRPFGPGW-----GDTV 353  
 Db 293 KNEKAALKVQEEYRKLGPLSFASINVLICFFLLVILWFSRDPGFMPLVAVVEGETK 352



Db 462 CTSNVATTTLFLPFPASMSRSIGINPLYIMLPCTLSASFAFMLPVATPPNAIVFTYGHK 521

Qy 528 VSDMARAGFLNIIIGVLAITLSINSWSIPFKLDTPPSWAH 568

Db 522 VADWVKGTGVIMNIIIGVFCVFLAVNTWGRAIFDLDFPDWAN 562

Search completed: June 30, 2005, 09:32:50  
Job time : 78.171 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:29:45 ; Search time 17.5314 Seconds  
(without alignments)  
3221.603 Million cell updates/sec

Title: US-10-017-479A-4  
Perfect score: 3058  
Sequence: 1 MATCHPALWAYRYLVILCL.....HSNTSQCLNPSNSTVPGGL 587

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2340.5	76.5	593	2 J46528	sodium/dicarbonyla
2	1622	53.0	568	2 I27911	Na+-coupled citrat
3	1259	41.2	595	2 A47714	Na+/sulfate cotran
4	1037.5	33.9	526	2 S43561	YCR37C homolog K08
5	1037.5	33.9	534	2 G88575	protein K08E5.2 (i
6	996.5	32.7	599	2 T21613	hypothetical prote
7	927	30.3	545	2 T18694	hypothetical prote
8	907.5	29.7	520	2 C89980	hypothetical prote
9	830.5	27.2	539	2 A88546	protein R107.1 (im
10	814.5	26.6	539	2 S30871	hypothetical prote
11	785	25.7	548	2 E71961	probable transport
12	781	25.5	552	2 F64546	sodium-dependent t
13	556	18.2	461	2 I64080	probable membrane
14	548	17.9	446	2 F70302	transporter (Pho87
15	525.5	17.2	462	2 B82510	transporter, Nadc
16	500.5	16.4	471	2 G81156	transporter, Nadc
17	496.5	16.2	456	2 AF0093	Sodium, sulfate sym
18	494.5	16.2	471	2 A81948	probable transmemb
19	478.5	15.6	443	2 C69205	sodium/dicarbonyla
20	468.5	15.3	487	2 C82215	probable transport
21	457	14.9	432	2 H64383	Na+ transporter -
22	455	14.9	478	2 A82336	probable transport
23	441	14.4	157	2 S36784	mucin - rat (fragm
24	383.5	12.5	517	2 C99840	conserved hypothet
25	344	11.2	867	2 T40336	probable MSF trans
26	313	10.2	881	2 S46633	probable membrane
27	305	10.0	513	2 I39534	hypothetical prote
28	305	10.0	894	2 S45135	probable membrane
29	294	9.6	612	2 S74936	sulfur deprivation

30	290	9.5	450	2	A10909	probable membrane
31	288	9.4	923	1	MMBY7C	probable membrane
32	256.5	8.4	487	1	B64795	ybd5 protein - Esc
33	256.5	8.4	487	2	G85560	probable membrane
34	256.5	8.4	487	2	C90710	probable membrane
35	242.5	7.9	589	2	H84072	sodium/sulfate sym
36	240.5	7.9	501	2	AE0578	citrate carrier [i
37	235.5	7.7	487	1	QEOCRS	ylgJ protein - Esc
38	234.5	7.7	487	2	B91122	hypothetical prote
39	234.5	7.7	487	2	A85967	hypothetical prote
40	228	7.5	424	2	F71205	hypothetical prote
41	223.5	7.3	608	2	AC0798	probable sodium/au
42	221	7.2	610	2	A85870	probable transport
43	221	7.2	610	2	H91025	probable transport
44	221	7.2	610	2	B65001	probable transport
45	220.5	7.2	610	2	AG0312	probable ion trans
46	213	7.0	425	2	C75019	transport protein
47	200.5	6.6	488	2	A81194	C4-dicarboxylate t
48	200.5	6.6	488	2	B81831	probable integral
49	200.5	6.6	610	2	H83165	probable sodium/au
50	197	6.4	618	2	F84409	arsenite transport
51	196.5	6.4	589	2	F75398	transporter, sodiu
52	188.5	6.2	484	2	AE0887	probable membrane
53	185	6.0	479	1	B64043	conserved hypothet
54	177.5	5.8	477	1	B64813	ybh1 protein - Esc
55	177.5	5.8	590	2	AG3106	transporter Atu447
56	177.5	5.8	590	2	E98180	transporter, sodiu
57	173.5	5.7	477	2	G85579	probable membrane
58	173.5	5.7	477	2	F90728	probable membrane
59	170.5	5.6	592	2	D95393	hypothetical prote
60	164.5	5.4	471	2	F71543	probable dicarboxy
61	160.5	5.2	476	2	B97096	2-oxoglutarate/mal
62	160	5.2	470	2	G86516	dicarboxylase tran
63	159	5.2	469	2	T35526	probable integral
64	157.5	5.2	482	2	F71969	hypothetical prote
65	157	5.1	472	2	F90078	hypothetical prote
66	155.5	5.1	585	2	S69216	sulfur deprivation
67	154	5.0	470	2	H72105	dicarboxylase tran
68	153	5.0	838	2	S28911	gene DN10 protein
69	148.5	4.9	475	2	A10010	probable membrane
70	147.5	4.8	411	2	G30154	arsenite transport
71	143	4.7	478	1	F69811	2-oxoglutarate/mal
72	142	4.6	436	2	S47723	arsenite efflux pu
73	139	4.5	429	2	F91175	arsenical pump mem
74	139	4.5	429	2	F86021	arsenical pump mem
75	135	4.4	513	2	A98839	Na+/H+-exchanging
76	135	4.4	513	2	G84864	Na+/H+-exchanging
77	134.5	4.4	425	2	F72315	conserved hypothet
78	134	4.4	513	2	A85697	Na+/H+ antiporter,
79	132.5	4.3	431	2	A11599	arsenic efflux pum
80	131.5	4.3	532	2	A57173	oculocutaneous alb
81	129	4.2	603	2	B59154	NADH2 dehydrogenas
82	129	4.2	603	2	T14026	NADH2 dehydrogenas
83	128.5	4.2	429	1	C41903	arsenical pump mem
84	128	4.2	515	2	E72089	ADP, ATP carrier p
85	126.5	4.1	429	1	B41902	arsenical pump mem
86	126.5	4.1	522	2	B64067	Na+/H+-exchanging
87	126	4.1	461	2	AB2185	hypothetical prote
88	125	4.1	430	2	G89982	hypothetical prote
89	125	4.1	515	2	B86534	ADP/ATP translocas
90	125	4.1	620	2	H82761	sulfur deprivation
91	124	4.1	614	2	AG0310	NADH2 dehydrogenas
92	123.5	4.0	548	2	AC3587	sulfur deprivation
93	122	4.0	429	2	G70528	probable arga prot
94	122	4.0	514	2	AG0723	regulator of intra
95	120.5	3.9	429	1	B25937	arsenical pump mem
96	118	3.9	603	1	DNHUN5	NADH2 dehydrogenas
97	116.5	3.8	370	2	AB0602	probable membrane
98	116.5	3.8	401	2	AF1800	transport protein
99	116	3.8	417	2	AF2992	arsenical pump mem
100	116	3.8	417	2	D98291	probable arsenical

ALIGNMENTS

RESULT 1  
146528  
sodium/dicarboxylate cotransporter - rabbit (fragment)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 14-Feb-1997 #sequence\_revision 14-Feb-1997 #text\_change 09-Jul-2004  
C:Accession: 146528  
R:Pajor, A.M.  
J. Biol. Chem. 270, 5779-5785, 1995  
A:Title: Sequence and functional characterization of a renal sodium/dicarboxylate cotransporter  
A:Reference number: 146528; MUID:95197598; PMID:7890707  
A:Accession: 146528  
A>Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-593 <P>  
A:Cross-references: UNIPROT:Q28615; EMBL:U12186; NID:g758383; PID:AAA99666.1; PID:g758383  
C:Superfamily: sodium/sulfate cotransporter

Query Match 76.5%; Score 2340.5; DB 2; Length 593;  
Best Local Similarity 74.0%; Pred. No. 1.6e-181;  
Matches 439; Conservative 71; Mismatches 74; Indels 9; Gaps 5;  
QY 1 MATCPALWAYRYLIVLCPLPILPLIVOTKEAYCAYSIIMALLMCTEALPLAVTA 60  
DB 1 MATCWOGLWAYRMVLLVPLLSLPLPLPRKEAYCAVAIIMLFWCTDALPLAVTA 60  
QY 61 LPPIVLPFLMGIMDASEVCIEYFKDTNLLFVGLMVAIAVEHNLHKLRIALQVLLIIGVR 120  
DB 61 LPLPCLCFPMGIMEASEVGLBYLKDNTNLFVGGLLLAIAVEHNLHKLRIALRVLLLTGVR 120  
QY 121 PALLLGLFMLVTAFLSMWISNTATTAMVPVPIGHVLEQLQSKDVEGGNNNPTFELOEE 180  
DB 121 PALLLGLFMMVTAFLSMWISNTASTAMVPVPIAHVLAQLNNTQSNVEGSDNPTFELOEP 180  
QY 181 CPQKEVTKL---DNG--QPUSA-PSEPTQKTQHHRFSGQLSICICYSASIGGIATLTG 234  
DB 181 SPQKETSVDKNGQAOPLPAVPLESGEHTQRLRFSGMSLCVCYSASIGGIATLTG 240  
QY 235 TPNLVLLQGVNSLFPQNGNVNFASFAGFAPPTMIILLALLMLOVLFLGVNFRKNFG 294  
DB 241 TPNLVLLQGVNSLFPQNGNVNFASFAGFAPPTMIILLALLMLOVLFLGVNFRKNFG 300  
QY 295 FGESEEE--RKOAAFOVIKTYRLLGPMSPFAEKTVTVLVLLVLMFTREPGFPGWGT 352  
DB 301 IREQEHEQQRQAARVVIQTYRLLGPMSPFAEKAFFILFVILVLLMFTREPGFPGWGNL 360  
QY 353 VPANEKQGSMSADGTVAIFSLVNFPIPSKIPGLMQDPKPKGKAPPAITLTKTWNDRM 412  
DB 361 VFDASGRVMVSDGSASILGVFLFVPSKIPGLTQDPDNPGRLLKAPPAITLTKTWNDRM 420  
QY 413 PWNIVLLGGGFALAKGSEGLSEWLGDKLTPLQHIPPSATAVILCLLAIETECTSNV 472  
DB 421 PWNIVLLGGGFALAKGSEGLSEWLGDKLTPLQHIPPSATAVILCLLAIETECTSNV 480  
QY 473 ATTTLLFLPILASMAQAICLHPLVYMLPCTLLAASLAFMLPVATPPNAIVFSGGLKYSDNA 532  
DB 481 ATTTLLFLPILASMAQAICLHPLVYMLPCTLLAASLAFMLPVATPPNAIVFSGGLKYSDNA 540  
QY 533 RAGFLNIIIGVLAITLSINWSPIFKLDTFPPSWAHNS-TSQCLLPNSNTPV 584  
DB 541 RAGFLNIIIGVLAITLSINWSPIFKLDTFPPSWAHNS-TTHCLASPPPTAPSP 593

RESULT 2  
JC7911  
Na+/sulfate cotransporter NaCT - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-2003 #sequence\_revision 31-Mar-2003 #text\_change 14-Jul-2003  
C:Accession: JC7911  
R:Inoue, K.; Zhuang, L.; Ganapathy, V.  
Biochem. Biophys. Res. Commun. 299, 465-471, 2002

A:Title: Human Na+-coupled citrate transporter: Primary structure, genomic organization,  
A:Reference number: JC7911; MUID:22334959; PMID:12445824  
A:Accession: JC7911  
A:Molecule type: mRNA  
A:Residues: 1-568 <INO>  
A:Cross-references: GB:AY151833  
C:Comment: This transporter classified as a tricarboxylate transporter represents the fil  
or the cellular entry of citrate by a process energized by the electrochemical Na+ gradient  
or the synthesis of fatty acid and chol 17p12-13  
C:Genetics:  
A:Gene: nact

Query Match 53.0%; Score 1622; DB 2; Length 568;  
Best Local Similarity 53.2%; Pred. No. 2.2e-123;  
Matches 309; Conservative 108; Mismatches 132; Indels 32; Gaps 6;  
QY 1 MATCPALWAYRYLIVLCPLPILPLIVOTKEAYCAYSIIMALLMCTEALPLAVTA 60  
DB 1 MASALSYVSKFSFVILFVTPLLPLVLMPEAKFVRCAYVILMAIYWCTEVPLAVTS 60  
QY 61 LPPIVLPFLMGIMDASEVCIEYFKDTNLLFVGLMVAIAVEHNLHKLRIALQVLLIIGVR 120  
DB 61 LMPVLLFPLFQILDSRQVCVQYMKDTNMLFLGLIVAVAVERNLHKLRIALTLLMVGA 120  
QY 121 PALLLGLFMLVTAFLSMWISNTATTAMVPVPIGHVLEQLQSKDVEGGNNNPTFELOEE 180  
DB 121 PARMLGFMGVTAFLSMWISNTATTAMVPVPIVEAILQMEATSAATEAG-----LELVDR 175  
QY 181 CPQKEVTKLNDGQPVSAPEBPTQKTQHHRFSGQLSICICYSASIGGIATLTGTPPNLV 240  
DB 176 GRAKE---LPGSQVIFEGPILGQEDQERKRLCKAMTLCICVAASTGGTATLTGTGPNV 232  
QY 241 LOGOVNSLFPQNGNVNFASFAGFAPPTMIILLALLMLOVLFLGVNFRKNFGEGEE 300  
DB 233 LIGQNNELPDSKDLVNFASFAPFAPNMLVMLFAWMLQVFNFRFKSGWGCGLESK 292  
QY 301 ERKQAAFOVIKTYRLLGPMSPFAEKTVTVLVLLVLMFTREPGFPGW-----GDTV 353  
DB 293 KNEKAALKVLEBYRKLGPLSPAEINVLICFPLVILFSDRPGFMPGMLTVAWVEGETK 352  
QY 354 PANEKQGSMSADGTVAIFSLVNFPIPSKIPGL-----MQDPKPKGKAPPAITLTKT 407  
DB 353 Y-----VSDATVAIFVATLFLFVPSQPKFNFRSQTEERKTP---FYPPPLLDWKV 401  
QY 408 VNDKMPNIVLLGGGFALAKGSEGLSEWLGDKLTPLQHIPPSATAVILCLLAIET 467  
DB 402 TOEKVPWGLVLLGGGFALAKGSEGLSVMGKQMEPLHVPAAITLILSLVAVFTE 461  
QY 468 CTSNVATTTLLPILASMAQAICLHPLVYMLPCTLLAASLAFMLPVATPPNAIVFSGGLK 527  
DB 462 CTSNVATTTLLPILASMAQAICLHPLVYMLPCTLLAASLAFMLPVATPPNAIVFSGGLK 521  
QY 528 VSDMARAGFLNIIIGVLAITLSINWSPIFKLDTFPPSWAH 568  
DB 522 VADMVKTGVIMNIIIGVFCVFLAVNTWGRAIFDLDFPDWAN 562

RESULT 3  
A47714  
Na+/sulfate cotransporter, renal - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Jun-1994 #sequence\_revision 02-Jun-1994 #text\_change 09-Jul-2004  
C:Accession: A47714  
R:Markovich, D.; Forgo, J.; Stange, G.; Biber, J.; Murer, H.  
Proc. Natl. Acad. Sci. U.S.A. 90, 8073-8077, 1993  
A:Title: Expression cloning of rat renal Na+/SO4(2-) cotransport.  
A:Reference number: A47714; MUID:93376745; PMID:7690140  
A:Accession: A47714  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-595 <MAB>  
A:Cross-references: UNIPROT:Q07782; GB:L19102; NID:g310182; PID:AAA41677.1; PID:g310183  
C:Superfamily: sodium/sulfate cotransporter

Query Match 41.2%; Score 1259; DB 2; Length 595;  
Best Local Similarity 42.3%; Pred. No. 5.7e-94;  
Matches 254; Conservative 123; Mismatches 176; Indels 48; Gaps 9;

QY 9 WAYRFLVILCLPIFLPLPLIVOTKEACVAYSIIIMALLWCTEALPLAVTALFIVLPP 68  
DB 8 FYVRFLVVFTVLLPLPLIRSKAEACVILFVIATFWITFTEALPLSITALLPLMFP 67  
QY 69 LMGIMDASEVCIEYFKDNTNIFVGLMVAIAVEHNLHKRIALQVLLIIGVRPALLLGF 128  
DB 68 MEGIMSSHTVASAYPKDFHLLIGVICLSTSEKNLHKRIALRMVMGVNPNALTLGF 127  
QY 129 MLVTAFLSMWISNTATTAMVPIGHAVLEQLQSGSKDVE-----GGNNNPTF 175  
DB 128 MSSTAFLSMWISNTSTAAMVPIEVAQAQITSAAEAEATQMTYFNESAQGLEVDETI 187  
QY 176 ELQECPCQKEVTK-----LDNGQPVASPSERTQ-----KTEHHRFSGLSLICY 222  
DB 188 IQQETNERKEKTKPALGSSNDKKG-VSSKMETEKNTVTGAKYRSKDHMMCKMLCIAY 246  
QY 223 SASIGGIATLTGTPNLVLOGQVNSLFPONGVNVNFASFPGFAFPTMIILLLLAWLQV 282  
DB 247 SSTIGLTIITGTSNLFSEHFNTRY-PDCRLNFGSFWLFPFVAVILLLLSWINLQW 305  
QY 283 LFLGVNFRKNGFGGEEERKQAAFOVIKTOYRLGLPMSFAEKTVTVLVLLVLMFTRE 342  
DB 306 LFLGVNFRKNGFGGEEERKQAAFOVIKTOYRLGLPMSFAEKTVTVLVLLVLMFTRE 342  
QY 343 PGFFPGWGTVPANKEGQSMASDGTVAIFISLVMFIIPSKIPGLMQDPKPKGLKAP--- 399  
DB 366 PGFVTGM-SVLFSYEPG--YVTDSTVALVAGILFLFLIPA-----KLTQKMTSTGDI 413  
QY 400 -----PAILTWTNDKMPNIVILGGGFGALAKGSEGLSEWLGDKLTPLQHIPPSAT 454  
DB 414 IAFDYSPLITWKEFGSFMPNDAILVGGFGALDQCQSGSSWIGSKLSPGSLPFWMLI 473  
QY 455 AVILCLIAITECTNSVATTLFLPILASMAQAICLHPLVYMLPCTLAASLAFMLPVAT 514  
DB 474 ILISSLIVTSLTEVASNPATITILPILSPLAEAHVNPHLHLLPSTLCTSFAPLLPVAN 533  
QY 515 PPNALVFGGLKVDMARAGFLNIIIGVLAITLSINSWSPIPKLDTFPPSWAHNSQTC 574  
DB 534 PPNALVFGGLKVIDMVKAGLVNGLVAVVLMGMFTWIBPMFNLHREYPSWAPDIVNQ 593  
QY 575 L 575  
DB 594 M 594

RESULT 4  
S43561  
YCR37C homolog K08E5.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 07-Sep-1994 #sequence\_revision 10-Nov-1995 #text\_change 24-Nov-1999  
C:Accession: S43561  
R:Kershaw, J.  
Submitted to the EMBL Data Library, March 1994  
A:Reference number: S43561  
A:Accession: S43561  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-526 <KER>  
A:Cross-references: EMBL:Z30974  
C:Genetics:  
A:Introns: 33/3; 122/2; 169/2; 214/2; 253/2; 481/3  
C:Superfamily: sodium/sulfate cotransporter

Query Match 33.9%; Score 1037.5; DB 2; Length 526;  
Best Local Similarity 40.9%; Pred. No. 4.2e-76;  
Matches 217; Conservative 108; Mismatches 156; Indels 49; Gaps 8;

QY 38 CAYSIIIMALLWCTEALPLAVTALFPIVLPPLMGIMDASEVCIEYFKDNTNIFVGLMVA 97

DB 37 CAYCVCIIVYMSEVMPLAVTAMLPVLPFLVGLVDANTTAKYEMNDTNLFIFGGLMA 96  
QY 98 IAVEHNLHKRIALQVLLIIGVRPALLLGLFMLVTAFLSMWISNTATTAMVPIGHAVLE 157  
DB 97 AAVEKCDLHERVALSVLRVCGSEPKWIMLGFWTMTALLSSFSISNTATTAMVPIQSVVQ 156  
QY 158 QLQSGKDKVEGGNNNPTFELQECPCQKEVTKLDNGQPVASPSERTQKTEHHRFSGLS 217  
DB 157 QL-----ISSFQHHPT-----NGERGLGCK-----KNATGLV 184  
QY 218 LCICYSASIGGIATLTGTPNLVLOGQVNSLFPONGVNVNFASFPGFAFPTMIILLLLAW 277  
DB 185 LSIICFAANIGTGATGTPSNLMLGQSLALFPKVDGSLNTVTWTFAYPLMLLCLFVAM 244  
QY 278 LMLQVFLGVNFRKNGFGGEEERKQAAFOVIKTOYRLGLPMSFAEKTVTVLVLLVVL 337  
DB 245 MTLVSFFL-----RDAPKDEAVTEMLKTRNNEPRMYAEKSVFVFCILLSL 293  
QY 338 WFTREPFGFPFGWGTVPANKEGQSMASDGTVAIFISLVMFIIPSKIPGLMQDPKPKGLK 397  
DB 294 WVRNPGVWPGFG-VFF-KKG--AYTDATSAMIVAFLLFVLPSERPDLATVIKKE-DLK 347  
QY 398 APPAILTWTNDKMPNIVILGGGFGALAKGSEGLSEWLGDKLTPLQHIPPSATVI 457  
DB 348 KRGCLMDWKMTQETFPMSVLLGGGFALAGVKESGLSLIGNSLSSIEHLPLWLQL 407  
QY 458 LCLLIAITECTNSVATTLFLPILASMAQAICLHPLVYMLPCTLAASLAFMLPVATPPN 517  
DB 408 TMLTAMVITNCSVTVTASIFVPIVATLAQRAGHPPFTMLPTTLASSFAIFVGTTPN 467  
QY 518 AIVFSGGLKVDMARAGFLNIIIGVLAITLSINSWSPIPKLDTFPPSWA 567  
DB 468 AIVFSGMWKVSMDAPVGGIISLELLVLTVLYNNSIAYLTPLLEFPTWA 517

RESULT 5  
G88575  
protein K08E5.2 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C:Accession: G88575  
R:anonymous, The C. elegans Sequencing Consortium.  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_elegans/ and  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: G88575  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-534 <STO>  
A:Cross-references: GB:chr\_III; PIDN:CAA83225.1; PID:G3878357; GSPDB:GN00021; CESP:K08E5.  
C:Genetics:  
A:Gene: K08E5.2  
A:Map position: 3  
C:Superfamily: sodium/sulfate cotransporter

Query Match 33.9%; Score 1037.5; DB 2; Length 534;  
Best Local Similarity 40.9%; Pred. No. 4.3e-76;  
Matches 217; Conservative 108; Mismatches 156; Indels 49; Gaps 8;

QY 38 CAYSIIIMALLWCTEALPLAVTALFPIVLPPLMGIMDASEVCIEYFKDNTNIFVGLMVA 97  
DB 45 CAYCVCIIVYMSEVMPLAVTAMLPVLPFLVGLVDANTTAKYEMNDTNLFIFGGLMA 104  
QY 98 IAVEHNLHKRIALQVLLIIGVRPALLLGLFMLVTAFLSMWISNTATTAMVPIGHAVLE 157  
DB 105 AAVEKCDLHERVALSVLRVCGSEPKWIMLGFWTMTALLSSFSISNTATTAMVPIQSVVQ 164  
QY 158 QLQSGKDKVEGGNNNPTFELQECPCQKEVTKLDNGQPVASPSERTQKTEHHRFSGLS 217  
DB 165 QL-----ISSFQHHPT-----NGERGLGCK-----KNATGLV 192



Db 400 LACSLNSTSKMPFFVFMQIIISIVVMTFSTNSATASIFIPISFKMAEAVGAHPLYF 459  
Qy 497 MLPCTLAASLAFMLPVPATPPNAIVFSGGLKVSMDARAGFLNIIGVLAITLSINSWSIP 556  
Db 460 SIPTAIGPSFSLPWPATANAIVETKIRIMIDWVSCGVNIFCIAITAINMTWAPW 519  
Qy 557 IPKLDTPPSWA--HS-----NTSQC 574  
Db 520 LFNMGTYPDYALRHATNMTGNSQC 544

RESULT 8

C99980  
hypothetical protein SA1732 [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: C89980  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaishi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: C89980  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-520 <KUR>  
A:Cross-references: UNIPROT:Q99SX1; GB:BA000018; PID:g13701709; PIDN:BA843002.1; GSPDB:C  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: SA1732  
C:Superfamily: sodium/sulfate cotransporter

Query Match 29.7%; Score 907.5; DB 2; Length 520;  
Best Local Similarity 37.0%; Pred. No. 1.4e-65;  
Matches 198; Conservative 97; Mismatches 159; Indels 81; Gaps 10;  
Qy 22 IFLLPL----PLIVQTKAYCAYSIILMALLWCTEALPLAVTALPPIVLPFLMGMIMDASE 77  
Db 42 LFLLLPFPHPQDLPKGVVLAITLWATWITAIPIATSLPILVPLPGLHILTEQ 101  
Qy 78 VCIEVFKDTNLIIFVGLGMVAIAVEHNLHKLIALQVLLIIGVRPALLLGLFMLVTAFLSM 137  
Db 102 VSSEYENDIIFLFGGFIILAIAMERWNLHTRVALTIINLIGASTSKILLGFMVATGFLSM 161  
Qy 138 WLSNTATTAMVPIGHAVLEQLQSGKQDVEGNNPTPELOECQKVEKTLKNDQPVSA 197  
Db 162 FVSNAAVMIMIPILGAIKEAH-----DLOEANTNQ----- 194  
Qy 198 PSEPRQTQEHHRFSQGLSLCYSASITGGIATLTGTPNLVLOGQVNSLPQNGNVN 257  
Db 195 -----STQKEKSLVLAIGVAGTIGGLTIGTPPILILKGQWQH-----GHIS 241  
Qy 258 PASWFGFAPPTMIILLLAWLQVLFLGVNFRNFGFGEGEREKQAQFVKTQYRL 317  
Db 242 FAKNMIVGPTIVILGITLWLYR----VAFRHLKYLPGQ-----TLIKQLDEL 290  
Qy 318 GPMSEAKTIVTVLVLVWVMTREPGFPGMGDTVPANKEQO--SMASDGTVAIFSLV 375  
Db 291 GOMKYEKVVQITFVLASLLWITRE-----FLKKWEVTSVADGTIAIFISIL 339  
Qy 376 MFIIPSKIPGLMQDKPKGLKAPPAILTWTNDKMPNIVILGGGFALAKGSEQSL 435  
Db 340 LFIIPAK-----NTEKHRIIDWE--VAKELPWGVLLIPGGGLALAKGISEGL 386  
Qy 436 SEWLGDKLTPLOHIPPSTAVATVILCLLIAIFTECTSNVATTTLFPLILASMAQICLHPY 495  
Db 367 AKWLGEQLKSLNGVSPILVIVITFVLFLTEVTSNTATATMILPILATLSVAVGVHPLL 446  
Qy 496 VMLPCTLAASLAFMLPVPATPPNAIVFSGGLKVSMDARAGFLNIIGVLATLSI 550  
Db 447 LMAPAAMAANCAYMLPVGTTPNAITFGSGKISIKQMASVGFWNLISAIILV 501

RESULT 10

S30871  
hypothetical protein R107.1 - Caenorhabditis elegans

RESULT 9

A88546  
protein R107.1 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: A88546  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/c\_elegans/ and www.sanger.ac.uk/Projects/C\_elegans/  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: A88546  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-539 <STO>  
A:Cross-references: UNIPROT:P32739; GB:chr\_III; PIDN:CAA78468.1; PID:g3879033; GSPDB:GN01  
A:Note: predicted using GeneFinder  
C:Genetics:  
A:Gene: R107.1  
A:Map position: 3  
C:Superfamily: sodium/sulfate cotransporter

Query Match 27.2%; Score 830.5; DB 2; Length 539;  
Best Local Similarity 32.7%; Pred. No. 2.5e-59;  
Matches 190; Conservative 126; Mismatches 198; Indels 67; Gaps 12;  
Qy 15 LIVLCLPIFLPLPIVOTKEAYCAYSIILMALLWCTEALPLAVTALPPIVLPFLMGMIMD 74  
Db 13 LLLVLLGP--LVAVPLLPFGPEYRCLFSIIFLSTWIGEAFFPIGVTSLPPLALYPLIQIVP 70  
Qy 75 ASEVCIEYFKDTNLIIFVGLGMVAIAVEHNLHKLIALQVLLIIGVRPALLLGLFMLVTAFL 134  
Db 71 SKQISPVYFKDSIVFLMCTLIMAAVEATGLHRRALKLLTKVGAQK----- 118  
Qy 135 LSMWISNTATTAMVPIGHAVLEQLQSGKQDVEGNNPTPELOECQKVEKTLKNDQNP 194  
Db 119 VSPFVSDTACTALMCPATAVALLMSMSDAVQHLKEDHRKPK--PPDDATVAEKLRIIDWTP 177  
Qy 195 VSASEPRQTQEHHRFSQGLSLCYSASITGGIATLTGTPNLVLOGQVNSLPQNGN 254  
Db 178 QDA-----GFCKALILACAHASLIGGTAITITSTGPNLVFRENHKKRYPEQOV 224  
Qy 255 VVNFASWFGFAPPTMIILLLAWLQVLFLGVN--ERKNFGEFGEGER--KQAQFQVTKT 312  
Db 225 TMTYLQWVFAIPPMPFVYLASYIILVCYFMGPSTFARWFERPERSKEEHLKLEKNIQT 284  
Qy 313 QYRLIGPMSFAEKTIVTVLVLVWVMTREPGFPGMGDTVPANKEGQSMASDGTVAIFI 372  
Db 285 MYEDLDGVNGEKSVFVFFILLIGSWISRDGFTPGWGDL-----PHRNFISDSVSGVLI 340  
Qy 373 SLVNFIIIPSKIPGLMQDKPKGLKAPPAILTWTNDKMPNIVILGGGFALAKGSEQ 432  
Db 341 SCILFVWP-----KDPFDPIDMAP--ILKWTDMKSKFSWCTLLIGAGYAISEGVDK 391  
Qy 433 SGLSE-----WLGDKLTPLOHIPPSTAVATVILCLLIAIFTECTSNVATTTLFPLILA 483  
Db 392 SGLSRLISCGMKNFVFGMSLPLQ-----LVTVTIIVIMTEFASNVSTGSIPIPSL 443  
Qy 484 SMAQICLHPILVNLPTCLAAFLPMLPVPATPPNAIVFSGGLKVSMDARAGFLNIIGV 543  
Db 444 GVAESMGVHLYALPTTVACSFAPMLPISTFPNAVVDYTKVISVEMIVCGFLINACI 503  
Qy 544 LAITLSINSWSIPILFKLDTPPSWAHSNTSQCLNPSNSTVP 584  
Db 504 LITSLNMTWTYFISLNIFF-----ENIVISSENSSYP 537

C:Species: Caenorhabditis elegans  
C>Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000  
C:Accession: S30871  
R:Thomas, K.  
submitted to the EMBL Data Library, July 1992  
A:Reference number: S30871  
A:Accession: S30871  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-539 <THO>  
A:Cross-references: EMBL:Z14092  
C:Genetics:  
A:Genetic code: SGC4  
A:Introns: 30/3; 80/3; 119/2; 217/2; 251/3; 293/2; 390/3; 416/1; 445/3; 491/3  
C:Superfamily: sodium/sulfate cotransporter  
C:Keywords: transmembrane protein

Query Match      26.6%; Score 814.5; DB 2; Length 539;  
Best Local Similarity    31.7%; Pred. No. 4.9e-58;  
Matches 186; Conservative 128; Mismatches 195; Indels 77; Gaps 12;

Qy 15 LIVLCPLFLLPLLVOTKEAYCAYSIIMALLWCTEALPVAVTALFPVLFVLGMGMD 74  
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 13 LLVLGGP--LVAVPLLPGPEYSCLFSSMFLSTYWIGEARPIGVTSLPFLALYBILQVP 70  
  
Qy 75 ASEWCIEYKDTNITLFGGLVMVATAVHNNLKHEIALQVLLIIQVRPALLLLGPMLVTAF 134  
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 71 SKQISPVYFKDSMWLFMCITLMANAVATGURHSIALKLTKVGAKQP----- 118  
  
Qy 135 LSMNISNTATTAMVPIGHAVLEQLQGSKKDVEGGNNP-----TFLEQECPOKEYTKL 189  
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 119 VSFFVSDTACTALMCPTAVALLMSWSDAVOHLKHEDSKPKPPDDATVAEKLSMDDMT PQ 178  
  
Qy 190 DNGQPVSAESPRTQKTEHRPFSGQLSCLICYSASIGGIATLTGTTPNLVLQGVNSLF 249  
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 179 DAG-----FCKAULACAHASLGIGTAITSTGNLFFRENIHKSY 219  
  
Qy 250 PONGNVNFASFAGFAPPTMIILLAWLMIQVLFLGVN-FRNKFGEGEEER-KQA AF 307  
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 220 PEGQVTWTYLQMMVFAMPMPFVYLLASYIILVCYFMGPSIFARWFESPSKEEAHLKL IE 279  
  
Qy 308 QVIKTQRLGLGPMFAKTVTLVLLVLVWFTREPQGFPGWGDTV FANEKGQSMA SDGT 367  
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 280 KNIQTMYDELGDVSWGSKSVFVFILLIGWISRDPGTFPGWGDL----PHSNFMVS DS V 335  
  
Qy 368 VAIRISLVMFIPSKIPLGMODPKPKGLKAPPAILTWKTVDNDKMPNIVILLGGG PALA 427  
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 336 SGVLISCILFWP-----KDPFDIDPWAP--ILKWTDMKSFSWSCILLIGAGV AIS 386  
  
Qy 428 KGSEQSG LSE-----WLGDKLTLPLQHIPPSPATAVILCLLIAFECTSNVAITTLF 478  
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 387 EGVDKSGLSLISCGMKNI FVGMSLPQ-----LTVTTIIVMTEFASNVTSGS IF 438  
  
Qy 479 LPILASAQAICLHPVLMPLCTLAASIAFMLVPATPNNAIVSFPGCLKYSDMARAG FL 538  
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 439 IPISLGAESMGVHPLVLAUPTTVACSFAMFLPISTPPNAVVDYTKVISVEMMVCG FL 498  
  
Qy 539 NIIGVLAITLSINSWSPIFKLDTPFSPWAHSNTSQCLLNPSNSTVP 584  
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 499 NIACMLTSLNMNTWTYFI SLNIFF-----ENIVWSSENSSY P 537

RESULT 11  
E71961  
probable transporter - Helicobacter pylori (strain J99)  
C:Species: Helicobacter pylori  
A:Variety: strain J99  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: E71961  
R:Alm, R.A.; Ling, J.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;  
Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path-

A:Reference number: A71800; MUID:99120557; PMID:9923682  
A:Accession: E71961  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-548 <ARN>  
A:Cross-references: UNIPROT:Q9ZML8; GB:AE001458; GB:AE001439; NID:g4154713; PIDN:AAD05784  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: jhp0200  
C:Superfamily: sodium/sulfate cotransporter

Query Match 25.7%; Score 785; DB 2; Length 548;  
Best Local Similarity 34.7%; Pred. No. 1.2e-55;  
Matches 181; Conservative 100; Mismatches 172; Indels 68; Gaps 14;

Qy 43 ILMALLMCTEALPLAVTALFPFVLVLPFLMGINDASEVCIEYFKDNTNLFVGGGLMVAVEH 102  
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:::  
68 VLMGIMWNTEAIDLPA TAL PLV LFNVSVDQFSSVSASYASPIIFLFWGGFIALSMQK 127  
Qy 103 WNLHKRALQVLLIGVRPALLLGFM LVTAFLSNWISNTATAMVPICHAVLEQLQGS 162  
Db |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
128 WNLHTRIALSILIVGTSPRLIILGFMTATGFLSMWVSNTATAVMMFPFGMSVLQLV--- 184  
Qy 163 KKDVEGGNNPTFELOECBPQKEVTK-----LDNGQPVSAPSEPRTKTOEHHRFS 213  
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
185 AKLVGENASNAPFOKE----ETKHAHGGINSVIHKGKDAOVIOEXT--TIYRTNFS 237  
Qy 214 QGLSLTCYSASIGGIATLTGTTNMLVLOGVNSLPQNGNV-VNPASFPGFAFTPMIL 272  
Db |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
238 ICMLGAYSASISGLSGTLTGTPNALLAGMYKTAF---NIEIDPAQVMVFCTPLAFIM 293  
Qy 273 LLALWLMLQVLFLGVNFRKNFGEGEBERQAQFQVTKTVYRLLGPMSPAECTVTVLFV 332  
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
294 LILAWLLTYVFPKLKIEIPG---GKE-----VVKSELKKURLSQAEISVGVI 342  
Qy 333 LLVVLTFRBEPFFPGWG---DTVPANEKGSMASDGTVAIFISLVMFIIPSKIPGLMQD 389  
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
343 LASLGWIFLDT-ILKSWGIKIDKI-----DSVTAMGVSA LFLTPANHQG--- 386  
Qy 390 PKXPGKLA KPAPAILTWKTVDNMKPNNIVILGGG FALAKGSEQSLSEWLDKLTPLQHI 449  
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
387 -----DRLLDW-GVAKKLPWDVULLFGGGLALSAQFSKTLGSLMIGHLVSGFSHL 435  
Qy 450 PPSATAVILCLLIAIFTCTSNVATTTTLFILASMAQAI---CLHPLYVMLPCTILAASL 506  
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
436 PILFIIVMWTLVWFLEITSNTATAAFLPVIGGVAMGMGYESHQSLLLTIPVALSATC 495  
Qy 507 AFMLPVATPPNAIVSFSGGLKVSDMARAGFLINLIIGVLAIT 547  
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
496 AFMLPVATPPNAIAYSGYVKITDMIKAGLWLNLVGVVLIS 536

RESULT 12  
F64546  
sodium-dependent transporter - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C:Accession: F64546  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, S.; J.D.; Kelley, J.M.; Cooton, M.D.; Weldman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.; Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N.  
A:title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: F64546  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-552 <YOM>  
A:Cross-references: UNIPROT:O25003; GB:AE000541; GB:AE000511; NID:g2313299; PIDN:AAD07282  
C:Superfamily: sodium/sulfate cotransporter

Query Match 25.5%; Score 781; DB 2; Length 552:

[illegible]

D	b		:       :       :       :	-DREXDHNTYFVL-----	173
L	35	F	S L S M M S N T A A M M P L I A M G I L S Q L -	C I C V S A S I G G I A T L T G T P N L V L O G V N S L F P Q N G	253
P	V	S	A P S E R T K T Q B H R F S Q G L S L C I C V S A S I G G I A T L T G T P N L V L O G V N S L F P Q N G	--LGIAYSASIGGGTUGVGPNNIAVSNL-	202
N	V	N	P A S W F G F A P P T M I L L L L A W L Q V L F --	--LGVNFRKNFGGEGEERKQAAFOVI	310
T	F	S	D W L Y G L P I M I L L P L M I G I L Y I I P K P K L H N F E Q T F --	---ENIE----	247
K	T	O	R L L G P M S F A E K T V V L F V L V L N F T R E P G F G G D T V F A N E K G --	--QSMA-S-DGT	367
M	N	M	P R I --LTPIIFPVIALTW----	--IPSGKINPFIISGLLGLOKNIASFDSI	292
V	I	F	I S L V M F I I P S K I P G L M Q D P K P G K L K A P P A I L T W K T V N D K M P W N I V I L L G G G F A L A	427	
V	A	L	L A A I V I -----CSTGVASNKIQTSDNDWGMVFPGGLTLIS	331	
K	G	S	B Q S G L S E W G D K L T P L ---QHPPSATAVILCLLIAIF---	TECTSNVAATFLPLP	480
A	V	L	K D S G A K I L A D S I V E M I D Q H P ----	YLIGLLVAAPITLPLEPTTSNITASAALLVP	385
I	L	A	S N A Q A I C L H P L Y V M L P C T L A A S L A F M L P V A T P P N A I V S F G L K V S D M A R A G F L N I	540	
I	F	I	S A O S L G M P E I G L A I I I G I G A C A F M L P V A T P P N A I V S F G S G V K Q S E M V K V G F L L N L	445	
I	G	V	L A I -T L S I N S W	553	
V	C	V	V V I A T M G Y M F W	459	
 RESULT 14 F70302 transporter (Pho87 family) - Aquifex aeolicus C:Species: Aquifex aeolicus C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004 C Accession: F70302 R Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham V. Nature 392, 353-358, 1998 A>Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus A Reference number: A70300; MID:98196666; PMID:9537320 A Accession: F70302 A>Status: preliminary; nucleic acid sequence not shown; translation not shown A Molecule type: DNA A Residues: 1-446 <AO> A Cross-references: UNIPROT:O66449; GB:AEO00670; NID:g2982779; PIDN:AAC06409.1 A Experimental source: strain VP5 C Genetics: A Gene: trns C Superfamily: probable transporter MJ0672  Query Match                  17.9%; Score 548; DB 2; Length 446; Best Local Similarity        31.2%; Pred. No. 1.5e-36; Matches 157; Conservative 74; Mismatches 163; Indels 110; Gaps 14					
Q	y	42	I L M A L W C T E A L P L A V T A L P P I V L F P L M G I M D A S E V C I E Y F K D N I - L F V G L M V A I A V	100	
D	b	34	L L L A A I L M I T E A L P L V T A L I P V S G V L L G V F D V - K T A L S Y F A H P L I F L F G G F V L A V A L	92	
Q	y	101	E H W N L H K R I A L Q V L L I G T V R --P A L L I G F M L V T A F L S M W I S N T A T T A M V P I G H A V L E Q	158	
D	b	93	S K Y Q I D E Y I A H K I V S A Q G F L P S V F L L -M L A T S L S M W I S N T T T A M L P L A U G I L A G	150	
Q	y	159	I Q S K K D V E G G N N N P T F E L Q E C P Q E V T K L D N G Q P S A P S E R P T K T Q E H R F S Q G L S L	218	
D	b	151	V-----RETAREKVFPFVLL	165	
Q	y	219	C I C V S A S I G G I A T L T G T P N L V L O G V N S L F P Q N G V N V N F A S W F G F A P P T M I L L L L A W L	278	
D	b	166	G I A Y S A S V G I T U G V G P P N G I A A G I L G L S P -----DWLKFGIPEVLILPFLLF--	215	

```
QY 279 MLQVLFLGVNFRKFGGEBERKQAAFOVIKTOYRLGPMSPAEKTTVTLVLLVVLW 338
Db 216 --AIIFL--VFPT-----SDLKVERVQEK-----FETPQRVLLIFL----- 252
QY 339 FTREPGFFPGWGTVPANEKQSMASDGTVAIFISLVNFIIPSKIPGLMQDPKPKGLKA 398
Db 253 FTALAWIFSKIAPIFEVKK-----YFDTVALLAVLFI----- 289
QY 399 PPAIITWTKMKNVNIILGGGFALAKSEQSGLSWLDKLTPLQHIPTSPATAVIL 458
Db 290 --RLDMDRDVKEGVSWGTLILFGGIALSGIMKKTGTAKFISQELVDVLHGLTFFLFLT 347
QY 459 CIIIAIF-TECTSNVATTTFLPILASMAQAICLHPLVMLPCTTLAASLAFMLPVATPPN 517
Db 348 IVLFVIFLTELMSNTATTALAPILFSTAQMGKPEMLVIPAAVAASCAFMPLPVATPPN 407
QY 518 AIVFSFGGLKVSMDMARAGFLNII 541
Db 408 AIVYGTGYIKOSQMMRVGLILNIV 431

RESULT 15
B82510
transporter, NadC family VCA0025 [imported] - Vibrio cholerae (strain N16961 serogroup O
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: B82510
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B
I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; PMID:20406833; PMID:10952301
A:Accession: B82510
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-462 <HEI>
A:Cross-references: UNIPROT:O9KNE0; GB:AE004346; GB:AE003853; NID:g9657401; PIDN:AAF9593
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0025
A:Map position: 2
C:Superfamily: probable transporter MJ0672
```

```
Query Match 17.2%; Score 525.5; DB 2; Length 462;
Best Local Similarity 29.7%; Pred. No. 1e-34;
Matches 153; Conservative 74; Mismatches 163; Indels 125; Gaps 13;

QY 42 IILMALLMCTEALPLAVTALFPVILFPLMGIMDASEVCIBYFKDTNILFVGGLMVAIAVE 101
Db 52 LAFTAVLWLTALHVTVTAILVPVMVFFGIFETQAAIINNPFANSIIFLFLGGFALAAAMH 111
QY 102 HNNLHKRIALQVLLIIGVRPALLLLGFWLVTAFLSMISNTATTAMVPIGHAVLEOLQG 161
Db 112 HOGLDKVIADKVLAMAQGMKSVAVFMLFGVTALLSMISNTATAAMMLPLVLGVL----- 166
QY 162 SKKDVEGNNPPTFELQECQKEVTKLDNGQPVSPAPSEPRTKQTEHHRFSQGLSLCIC 221
Db 167 SKVDADKQKRSYTVFL-----LGVA 186
QY 222 YSASIGGIATITGTPNLVQGVNSLFPQNGVNVFASWFGFAPPTMIILLLLAMLWLQ 281
Db 187 YSASIGGIATLVGSPPPNAIAAEVG-----LSFTDMWKFLPTAMWMLPMA----- 232
QY 282 VLFGLVNFNFKFGGEBERKQAAFOVIKTOYRLGPMSPAEKTTVTLVLLVWLWFR 341
Db 233 -----IAILYFLKP-----TLNGMPFL-----DR 252
QY 342 EPGFFPGWGTVPANEKQSMASDGTVAIF-ISLWFIIPSKIPGLMQDPKPKGLKAPP 400
Db 253 AP---VNW-----DKQKV-----TLGIFGLTVFLVIFSPINAALGGFKSFDTLVALG 298
QY 401 AILT-----WKTVNDKMPWNITVILGGGFALAKSEQSGLSWLDKLTPL-QHIP 451
```

```
Db 299 AIIMLSFARVVHWEIKQTADWGVLLFGGLCLSNVLKQTGTSVFLANALSDMVSHMGI 358
QY 452 SATAVILCLLIIAIFTCTSNVATTTTLPLILASMAQAICLHPLVYMLPCTTLAASLAFMLP 511
Db 359 FVIVILVATFVVFLTEFASNTASAALLIPVFATVAEAFGMSPVLLSVLIAVAASCAFMPLP 418
QY 512 VATPPNAIVFSFGGLKVSMDMARAGFLNII--TGVL 544
Db 419 VATPPNAIVFASGHIKOSEKMRVGLYINIACTIGLL 453
```

Search completed: June 30, 2005, 09:00:17  
Job time : 19.5314 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:39 ; Search time 78.9921 Seconds  
(without alignments)  
3805.323 Million cell updates/sec

Title: US-10-017-479A-4

Perfect score: 3058

Sequence: 1 MATCPALWAYRFYLIVLCL.....HSNTSQCLLNPSNVTVPGL 587

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3058	100.0	587	2 Q35055	O35055 rattus norv
2	2956	96.7	587	1 S132 RAT	P70545 rattus norv
3	2837	92.8	586	1 S132 MOUSE	Q9es88 mus musculu
4	2392	78.2	592	1 S132_HUMAN	Q13183 homo sapien
5	2340.5	76.5	593	1 S132_RABIT	Q28615 oryctolagus
6	2178	71.2	605	2 Q866R0	Q866R0 didelphis m
7	1971.5	64.5	622	2 O57661	O57661 xenopus lae
8	1967.5	64.3	624	2 Q642N5	Q642N5 xenopus tro
9	1867	61.1	613	2 Q6P3L2	Q6P3L2 brachydanio
10	1862	60.9	613	2 Q803K7	Q803K7 brachydanio
11	1768.5	57.8	586	2 Q6AZR9	Q6AZR9 xenopus lae
12	1624	53.1	568	2 Q86YT5	Q86YT5 homo sapien
13	1610	52.6	568	2 Q6ZMG1	Q6ZMG1 homo sapien
14	1536.5	50.2	572	2 Q67BT3	Q67BT3 mus musculu
15	1513.5	49.5	572	2 Q8CJ44	Q8CJ44 rattus norv
16	1378.5	45.1	600	1 S133 MOUSE	Q91y63 mus musculu
17	1377.5	45.0	600	1 S133 RAT	Q920z5 rattus norv
18	1364.5	44.7	602	2 Q6EJG8	Q6EJG8 xenopus lae
19	1364.5	44.6	602	2 Q6DJN7	Q6DJN7 xenopus lae
20	1355	44.3	602	1 S133_HUMAN	Q8wvt9 homo sapien
21	1354	44.3	602	2 Q81VB1	Q81VB1 homo sapien
22	1325	43.3	595	1 S131_HUMAN	Q9bz22 homo sapien
23	1322	43.2	601	2 Q9W7I2	Q9W7I2 pseudopiear
24	1316.5	43.1	583	2 Q6NY28	Q6NY28 brachydanio
25	1316.5	43.1	583	2 Q6PE27	Q6PE27 brachydanio
26	1311.5	42.9	596	2 Q6NWY4	Q6NWY4 brachydanio
27	1259	41.2	595	1 S131 RAT	Q07782 rattus norv
28	1256	41.1	627	1 S134_HUMAN	Q9uk94 homo sapien
29	1253.5	41.0	626	2 Q8N631	Q8N631 homo sapien
30	1253	41.0	601	2 Q7ZWL2	Q7ZWL2 xenopus lae
31	1248	40.8	625	2 Q8BZ82	Q8BZ82 mus musculu

ALIGNMENTS

32	1236	40.4	594	1 S131_MOUSE	Q9jh14 mus musculu
33	1234	40.4	594	2 Q80YB5	Q80YB5 mus musculu
34	1228	40.2	619	2 Q6L970	Q6L970 anguilla ja
35	1184.5	38.7	520	2 Q8N8K4	Q8N8K4 homo sapien
36	1043	34.1	566	1 NAD3_CAEEL	Q21339 caenorhabdi
37	1042.5	34.1	572	1 IND1_DROME	Q9vvt2 drosophila
38	1042.5	34.1	590	2 Q7KUS6	Q7KUS6 drosophila
39	1007	32.9	582	1 NAD1_CAEEL	Q73655 caenorhabdi
40	968	31.7	896	2 Q7Q1T2	Q7q1t2 anopheles g
41	937	30.6	552	2 Q8ENE3	Q8ene3 oceanobacil
42	930	30.4	450	2 Q75MH3	Q75mh3 homo sapien
43	925	30.2	577	1 YK56_CAEEL	P45556 caenorhabdi
44	907.5	29.7	520	2 Q99SX1	Q99sx1 staphylococ
45	907.5	29.7	520	2 Q7A4P8	Q7a4p8 staphylococ
46	907.5	29.7	520	2 Q6GFE0	Q6gfe0 staphylococ
47	906.5	29.6	520	2 Q8NV55	Q8nvs5 staphylococ
48	906.5	29.6	520	2 Q6G816	Q6g816 staphylococ
49	894	29.2	504	2 Q86B89	Q86b89 drosophila
50	889.5	29.1	551	1 NAD2_CAEEL	P32739 caenorhabdi
51	877	28.7	302	2 Q68D44	Q68d44 homo sapien
52	873.5	28.6	505	2 Q811B6	Q811b6 drosophila
53	871	28.5	562	1 IND2_DROME	Q9vdc0 drosophila
54	866.5	28.3	546	2 Q65NG0	Q65nc0 bacillus li
55	831.5	27.2	540	2 Q8LGR8	Q8lgr8 arabidopsis
56	829.5	27.1	540	2 Q9MAW4	Q9maw4 arabidopsis
57	826	27.0	462	2 Q9FEG1	Q9feg1 arabidopsis
58	820.5	26.8	540	2 Q9FEH9	Q9feh9 arabidopsis
59	794	26.0	626	2 Q6AQ77	Q6aq77 deauletale
60	792	25.9	548	2 Q6ZPH7	Q6zph7 oryza sativ
61	785	25.7	548	2 Q9ZML8	Q9zml8 helicobacte
62	781	25.5	552	2 Q25003	Q25003 helicobacte
63	775	25.3	524	2 Q6NKS9	Q6nks9 corynebacte
64	773	25.3	510	2 Q8NTS7	Q8nts7 corynebacte
65	773	25.3	527	2 Q6M8D0	Q6m8d0 corynebacte
66	763.5	25.0	495	2 Q67T81	Q67t81 symbiobacte
67	755	24.7	474	2 Q65MM0	Q65mm0 bacillus li
68	754	24.7	587	2 Q8FU27	Q8fu27 corynebacte
69	718.5	23.5	510	2 Q7UUK9	Q7uuk9 rhodopirell
70	714.5	23.4	414	2 Q7PF67	Q7pf67 anopheles g
71	659.5	21.6	440	2 Q81V67	Q81v67 homo sapien
72	644	21.1	450	2 Q8FM25	Q8fm25 corynebacte
73	613	20.0	334	2 Q6ZNH7	Q6znh7 homo sapien
74	565	18.5	464	2 Q8VTX1	Q8vtx1 leptospira
75	565	18.5	464	2 Q9AED9	Q9aed9 leptospira
76	565	18.5	464	2 Q72QJ5	Q72qj5 leptospira
77	561	18.3	464	2 Q8F5L4	Q8f5l4 leptospira
78	559	18.3	464	2 Q984F3	Q984f3 leptospira
79	557.5	18.2	470	2 Q5SR88	Q5sr88 manheimia
80	556	18.2	461	1 Y608_HABIN	Y57486 haemophilus
81	548	17.9	446	2 Q66449	Q66449 aquifex aeo
82	545.5	17.8	463	2 Q9CLQ0	Q9clq0 pasteurella
83	537	17.6	451	2 Q7M8Y1	Q7m8y1 wolinnella s
84	535	17.5	464	2 Q7ZGJ3	Q7zgj3 leptospira
85	525.5	17.2	462	2 Q9KNE0	Q9kne0 vibrio chol
86	519.5	17.0	456	2 Q87Q94	Q87q94 vibrio para
87	518	16.9	463	2 Q65SS5	Q65ses manheimia
88	513.5	16.8	456	2 Q6LQW4	Q6lqm4 photobacter
89	506.5	16.6	456	2 Q7MLG6	Q7mlg6 vibrio vuln
90	506.5	16.6	456	2 Q8D908	Q8d908 vibrio vuln
91	500.5	16.4	471	2 Q9K032	Q9k032 neisseria m
92	498	16.3	459	2 Q66615	Q66615 yersinia ps
93	496.5	16.2	456	2 Q8ZHX0	Q8zhx0 yersinia pe
94	495	16.2	459	2 Q8CZT3	Q8czt3 yersinia pe
95	494.5	16.2	471	2 Q9JVA3	Q9jva3 neisseria m
96	492	16.1	461	2 Q7N822	Q7n822 photorhabdu
97	490	16.0	474	2 Q8GF70	Q8gf70 photorhabdu
98	478.5	15.6	443	2 Q26881	Q26881 methanobact
99	468.5	15.3	487	2 Q9KSE3	Q9kse3 vibrio chol
100	464.5	15.2	471	2 Q6LSU9	Q6lsu9 photobacter

```

RESULT 1
C35055 PRELIMINARY; PRT; 587 AA.
AC 035055;
DT 01-JAN-1998 (TREMELrel. 05, Created)
DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Sodium/dicarboxylate cotransporter 1 (NA(+)/dicarboxylate
DE cotransporter 1) (Kidney dicarboxylate transporter) (SDCTL) (Organic
DE anion transporter 1) (OAT1).
GN Name=Nadc-1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=98355911; PubMed=9691021;
RA Sekine T., Cha S.H., Hosoyamada M., Kanai Y., Watanabe N., Furuta Y.,
RA Fukuda K., Igarashi T., Endou H.;
RT "Cloning, functional characterization, and localization of a rat renal
RT Na+-dicarboxylate transporter.";
RL Am. J. Physiol. 275:F298-F305(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=97373539; PubMed=9228014; DOI=10.1074/jbc.273.30.18526;
RA Sekine T., Watanabe N., Hosoyamada M., Kanai Y., Endou H.;
RT "Expression cloning and characterization of a novel multispecific
RT organic anion transporter.";
RL J. Biol. Chem. 272:18526-18529(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=98361968; PubMed=9694847; DOI=10.1074/jbc.273.33.20972;
RA Chen X.Z., Shayakul C., Berger U.V., Tian W., Hediger M.A.;
RT "Characterization of a rat Na+-dicarboxylate cotransporter.";
RL J. Biol. Chem. 273:20972-20981(1998).
CC -1- FUNCTION: MAY MEDIATE ELECTROGENIC, SODIUM-DEPENDENT TRANSPORT OF
CC MOST KREBS CYCLE INTERMEDIATES, INCLUDING CITRATE, SUCCINATE,
CC ALPHA-KETOGLUTARATE, AND OXALOACETATE. TRANSPORT OF CITRATE WAS PH
CC SENSITIVE. IT MAY HAVE A SODIUM-SUBSTRATE STOICHIOMETRY OF 3:1. AT
CC LEAST ONE NA(+) SEEMS TO BIND TO THE TRANSPORTER BEFORE THE
CC SUBSTRATE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN S3 SEGMENTS OF KIDNEY PROXIMAL
CC TUBULES AND IN ENTEROCYTES LINING THE INTESTINAL VILLI. ALSO
CC DETECTED IN LUNG BRONCHIOLI, THE EPIDIDYMIS, AND LIVER.
DR EMBL; AB001321; BAA28609.1; -.
DR EMBL; AF058714; AAC31165.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015293; F:symporter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001898; Na/sul_symport.
DR Pfam; PF00939; Na_sulph_sym; 1.
DR TIGRFAMs; TIGR00785; dss; 1.
DR PROSITE; PS01271; NA_SULFATE; 1.
KW Sodium transport; Symport; Transmembrane; Transport.
FT TRANSMEM 13 33 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT TRANSMEM 114 134 POTENTIAL.
FT DOMAIN 123 126 POLY-LEU.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT DOMAIN 272 275 POTENTIAL.
FT TRANSMEM 319 339 POTENTIAL.
FT TRANSMEM 367 387 POTENTIAL.
FT TRANSMEM 445 465 POTENTIAL.

```

---

```

FT TRANSMEM 477 497 POTENTIAL.
FT TRANSMEM 506 526 POTENTIAL.
FT TRANSMEM 535 555 POTENTIAL.
FT CARBOHYD 570 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 587 AA; 64142 MW; 844847C7F8F9CDA CRC64;

Query Match 100.0%; Score 3058; DB 2; Length 587;
Best Local Similarity 100.0%; Pred. No. 3.4e-223; Indels 0; Gaps 0;
Matches 587; Conservative 0; Mismatches 0;

Qy 1 MATCWPAALWAYFYLVLCPIFLPLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60
Db 1 MATCWPAALWAYFYLVLCPIFLPLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60
Qy 61 LPPVILFPLMGINDASEVCIEYFKDNIILFVGLVMVAIAVEHWNHLKRIALQVLLIIGVR 120
Db 61 LPPVILFPLMGINDASEVCIEYFKDNIILFVGLVMVAIAVEHWNHLKRIALQVLLIIGVR 120
Qy 121 PALLLLGFMVLTAFLSMTISNTATTAMVPIGHAVLEQLQSGKDVGGNNNPTEFLOEE 180
Db 121 PALLLLGFMVLTAFLSMTISNTATTAMVPIGHAVLEQLQSGKDVGGNNNPTEFLOEE 180
Qy 181 CPQKVTKLDNGQPVSAPEPRTKTQEHRRFSQGLSLCICYSASIGGIATLTGTTPLNV 240
Db 181 CPQKVTKLDNGQPVSAPEPRTKTQEHRRFSQGLSLCICYSASIGGIATLTGTTPLNV 240
Qy 241 LQGVNSLFPQNGVNVNPFASWFGFAPPTMIILLLAWLWLVFLGVNFRKNGFGEGEE 300
Db 241 LQGVNSLFPQNGVNVNPFASWFGFAPPTMIILLLAWLWLVFLGVNFRKNGFGEGEE 300
Qy 301 ERKQAAFOVTKQYRLGLGPMFAEKTIVLVLLVVLVLTRETPGPPGMDTVFANEKGQ 360
Db 301 ERKQAAFOVTKQYRLGLGPMFAEKTIVLVLLVVLVLTRETPGPPGMDTVFANEKGQ 360
Qy 361 SMASDGTVAIFISLVNFIIPSKIPLGMQDPKPKGLKAPPAILTWKTVNDKMPNIVILL 420
Db 361 SMASDGTVAIFISLVNFIIPSKIPLGMQDPKPKGLKAPPAILTWKTVNDKMPNIVILL 420
Qy 421 GGGFALAKSESGSLSEWLGDKLTPLQHPPTATAVILCLLTAIFTECTSNVATTTFLP 480
Db 421 GGGFALAKSESGSLSEWLGDKLTPLQHPPTATAVILCLLTAIFTECTSNVATTTFLP 480
Qy 481 ILASMAQALCLHPLYVNLPTCLLAASLAFMLPVATPPNAIVFSGGLKVDMARAGFLNI 540
Db 481 ILASMAQALCLHPLYVNLPTCLLAASLAFMLPVATPPNAIVFSGGLKVDMARAGFLNI 540
Qy 541 IGVLAITLSINSWSIPFKLDTFPPSWAHGNTSQCLNPSNSTVPGGL 587
Db 541 IGVLAITLSINSWSIPFKLDTFPPSWAHGNTSQCLNPSNSTVPGGL 587

RESULT 2
S132_RAT
ID S132_RAT STANDARD; PRT; 587 AA.
AC P70545;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Solute carrier family 13, member 2 (Intestinal sodium/dicarboxylate
DE cotransporter) (Na(+)/dicarboxylate cotransporter 1) (NADC-1).
GN Name=Slc13a2; Synonyms=Nadcl, Sdctl;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RX MEDLINE=97107437; PubMed=8950177; DOI=10.1016/S0167-4781(96)00138-8;
RA Khatiri I.A., Kovacs S.V.B., Forstner J.F.;
RT "Cloning of the cDNA for a rat intestinal Na+/dicarboxylate
RT cotransporter reveals partial sequence homology with a rat intestinal

```

RT mucin.";  
 RN [2]  
 RA Biochim. Biophys. Acta 1309:58-62(1996).  
 RP CONCEP.TM TRANSLATION.  
 RL Unpublished observations (SBP-2002).  
 CC -!- FUNCTION: Cotransport of sodium ions and dicarboxylates such as  
 CC succinate and citrate.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: Belongs to the SLC13A transporter family. NADC  
 CC subfamily.  
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to the  
 CC introduction of a number of frameshifts to maximize the similarity  
 CC with orthologs from other species.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U51153; AAB97095.1; ALT\_FRAME.  
 DR EGD; 61920; SLC13A2.  
 DR InterPro: IPR001898; Na/sul\_sympot.  
 DR Pfam: PF00939; Na sulph\_symp; 1.  
 DR TIGRFAMs: TIGR00785; daga; 1.  
 DR PROSITE: PS01271; NA\_SULFATE; 1.  
 KW Sodium transport; Symport; Transmembrane; Transport.  
 FT TRANSMEM 13 33 Potential.  
 FT TRANSMEM 53 73 Potential.  
 FT TRANSMEM 85 105 Potential.  
 FT TRANSMEM 135 155 Potential.  
 FT TRANSMEM 263 283 Potential.  
 FT TRANSMEM 329 349 Potential.  
 FT TRANSMEM 367 387 Potential.  
 FT TRANSMEM 407 427 Potential.  
 FT TRANSMEM 445 465 Potential.  
 FT TRANSMEM 477 497 Potential.  
 FT TRANSMEM 506 526 Potential.  
 FT TRANSMEM 535 555 Potential.  
 SQ SEQUENCE 587 AA; 64206 MW; 964D80661F1EBE0A CRC64;  
 Query Match 96.7%; Score 2956; DB 1; Length 587;  
 Best Local Similarity 97.4%; Pred. No. 1.8e-215;  
 Matches 573; Conservative 2; Mismatches 11; Indels 2; Gaps 2;  
 QY 1 MATCHPALWAYRFLYLIVLCPLIFLPLPLIVQTKAYCAYSILMALLWCTEALPLAVTA 60.  
 DB 1 MATCHPALWAYRFLYLIVLCPLIFLPLPLIVQTKAYCAYSILMALLWCTEALPLAVTV 60  
 QY 61 LPPVILFPLMGIMDASEVICIEFKDTNLFVGLMVATAVEHNLHKKRIALQVLLIIGVR 120  
 DB 61 LPPVILFPLMGIMDASE-GLEYFKDTNLFVGLMVATAVEHNLHKKRIALQVLLIIGVR 119  
 QY 121 PALLLGLFMLVAFSLMISNTATTAMVPIGHAVLEOLQGSKDVGGNNNPTPELOBE 180  
 DB 120 PALLLGLFMLVAFSLMISNTATTAMVPIGHAVLEOLQGSKDVGGNNNPTPELOBE 179  
 QY 181 CPQKEVTKLDNGQPVSAPEPRKTQTBHHRFSQGLSLCICYSASIGIATLTGTPNLV 240  
 DB 180 CPQKEVTKLDNGQPVSAPEPRKTQTBHHRFSQGLSLCICYSASIGIATLTGTPNLV 239  
 QY 241 LOGQVNSLFPQNGNTNPNASWFGAFPTMIILLLAWLQVLFVGNFRKNFGEGEE 300  
 DB 240 LOGQVNSLFPQNGNTNPNASWFGAFPTMIILLLAWLQVLFVGNFRKNFGEGEE 299  
 QY 301 ERKQAAFOVIKTOYRLGLGPMSPAEKTV-TVLVPLVLVLMWTFEPGFFPGWGTDFVANSKG 359  
 DB 300 ERKQAAFOVIKTOYRLGLGPMSPAEKTV-TVLVPLVLVLMWTFEPGFFPGWGTDFVANSKG 359  
 QY 360 QSMASDGTVAIFISLWMFIIPSKIPGLMEDPKPKLAPPAILTWKTVNDKMPNIVIL 419

DB 360 QSMPSDGTVAIFISLWMFIIPSKIPGLMEDPKPKLAPPAILTWKTVNDKMPNIVIL 419  
 QY 420 LGGGFALAKGSEQSGLSWLGDKLTPLOHIPPSSATAVILCLLIAIFTECTSNVATITLFL 479  
 DB 420 LGGGFALAKGSEQSGLSWLGDKLTPLOHIPPSSATAVILCLLIAIFTECTSNVATITLFL 479  
 QY 480 PILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVPSFGGLKVDMARAGFLIN 539  
 DB 480 PILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVPSFGGLKVDMARAGFLIN 539  
 QY 540 IIGVLATLTSNSWSIPKIDTPPSWAHSNTSOCLNPSNSTVPGGL 587  
 DB 540 IIGVLATLTSNSWSIPKIDTPPSWAHSNTSOCLNPSNSTVPGGL 587  
 RESULT 3  
 S132\_MOUSE  
 ID S132\_MOUSE STANDARD; PRT; 586 AA.  
 AC Q9ES88;  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Solute carrier family 13, member 2 (Renal sodium/dicarboxylate  
 DE cotransporter) (Na(+)/dicarboxylate cotransporter 1) (NADC-1).  
 GN Name=SLC13a2; Synonyms=Nadcl, Sdct1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=20425391; PubMed=10965927;  
 RA Pajor A.M., Sun N.N.;  
 RT "Molecular cloning, chromosomal organization and functional  
 RT characterization of a sodium/dicarboxylate cotransporter from mouse  
 RT kidney.";  
 RL Am. J. Physiol. 279:F482-F490(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raba S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Cotransport of sodium ions and dicarboxylates such as  
 CC succinate and citrate.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- TISSUE SPECIFICITY: Highly expressed in kidney and small  
 CC intestine. Not detectable in brain, heart, stomach and skeletal  
 CC muscle.  
 CC -!- SIMILARITY: Belongs to the SLC13A transporter family. NADC  
 CC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; AF201903; AAG15426.1; -;  
 DR EMBL; BC013493; AAH13493.1; -;  
 DR MGD; MG1.1276558; SLC13A2.  
 DR InterPro; IPR001898; Na/sul\_sympoort.  
 DR Pfam; PF00939; Na\_sulph\_symp; 1.  
 DR TIGRFAMs; TIGR00785; daas; 1.  
 DR PROSITE; PS01271; NA\_SULFATE; 1.  
 KW Sodium transport; Symport; Transmembrane; Transport.  
 FT TRANSMEM 13 33 Potential.  
 FT TRANSMEM 53 73 Potential.  
 FT TRANSMEM 86 106 Potential.  
 FT TRANSMEM 215 235 Potential.  
 FT TRANSMEM 264 284 Potential.  
 FT TRANSMEM 319 339 Potential.  
 FT TRANSMEM 366 386 Potential.  
 FT TRANSMEM 407 427 Potential.  
 FT TRANSMEM 445 465 Potential.  
 FT TRANSMEM 478 498 Potential.  
 FT TRANSMEM 506 526 Potential.  
 FT TRANSMEM 535 555 Potential.  
 SQ SEQUENCE 586 AA; 64110 MW; 56EAE21532833EE CRC64;

Query Match 92.8%; Score 2837; DB 1; Length 586;  
 Best Local Similarity 92.1%; Pred. No. 1.9e-206;  
 Matches 539; Conservative 25; Mismatches 21; Indels 0; Gaps 0;

Qy 1 MATCPALWAYRFLVILVLCPIFLPLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60  
 Db 1 MATCPALWAYRFLVILVLCPIFLPLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60  
 Qy 61 LFPVLVPLMGIMDASEVCIEYFDKTNILFVGLWVAIAVEHNLKRIALQVLLIIGVR 120  
 Db 61 LFPVLVPLMGIMDASEVCIEYFDKTNILFVGLWVAIAVEHNLKRIALQVLLIIGVR 120  
 Qy 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLOGSKDKVGGNNPTFELQEE 180  
 Db 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLOGSKDKVGGNNPTFELQEE 180  
 Qy 181 CPQEVTKLDNGQVSPASBPRTOKTOEHRFSSQGLSLCYASIGGIATLTGTTPLV 240  
 Db 181 SPQKEETKLDNGQVSPASBPRTOKTOEHRFSSQGLSLCYASIGGIATLTGTTPLV 240  
 Qy 241 LOGQVNSLFPONGVNVFASFVGFAPFTMTILLLAWLQVLPLGVNFRKNFGFGE 300  
 Db 241 LOGQVNSLFPONGVNVFASFVGFAPFTMTILLLAWLQVLPLGVNFRKNFGFGE 300  
 Qy 301 ERKQAFQVKTQVRLIGPMSFAEKTVTTLVFLVILVWFTREPGFFPGWGDVPA 360  
 Db 301 ERKQAFQVKTQVRLIGPMSFAEKTVTTLVFLVILVWFTREPGFFPGWGDVPA 360  
 Qy 361 SMASDGTVAIFISIVMTIISKIPGLMOPDKPKGLKAPAILTWKTNKMPNIVILL 420  
 Db 361 SMASDGTVAIFISIVMTIISKIPGLMOPDKPKGLKAPAILTWKTNKMPNIVILL 420  
 Qy 421 GGGFALAKGSEGLSEWLGDKLTPLOHIPPASATAVILCLLIAIFTECTSNVATTLFLP 480  
 Db 421 GGGFALAKGSEGLSEWLGDKLTPLOHIPPASATAVILCLLIAIFTECTSNVATTLFLP 480  
 Qy 481 ILASMAQAICLHPLVYVLMPTCTLAASAFMLPVPATPPNAIVFSGGLKVSMDARAGFLINI 540  
 Db 481 ILASMAQAICLHPLVYVLMPTCTLAASAFMLPVPATPPNAIVFSGGLKVSMDARAGFLINI 540  
 Qy 541 IGVLAITLSNSWSIPFKLDTFSPSWAHSNTSQCLNPSNTVPG 585  
 Db 541 IGVLAITLSNSWSIPFKLDTFSPSWAHSNTSQCLNPSNTVPG 585

RESULT 4  
 S132 HUMAN STANDARD; PRT; 592 AA.  
 AC Q13183; 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Solute carrier family 13, member 2 (Renal sodium/dicarboxylate cotransporter) (Na(+)/dicarboxylate cotransporter 1) (NaDC-1).  
 DE Name=SLC13A2; Synonyms=NADC1, SDC11;  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=96199379; PubMed=8967342;  
 RA Pajor A.M.;  
 RT "Molecular cloning and functional expression of a sodium-dicarboxylate cotransporter from human kidney";  
 RL Am. J. Physiol. 270:F642-F648(1996).  
 CC -!- FUNCTION: Cotransport of sodium ions and dicarboxylates such as succinate and citrate.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: Belongs to the SLC13A transporter family. NADC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U26209; AAA98504.1; -;  
 CC Genew; HGNC:10917; SLC13A2.  
 CC MIM; 604148; -;  
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.  
 CC GO; GO:0005624; C:membrane fraction; TAS.  
 CC GO; GO:0015361; F:flow affinity sodium:dicarboxylate symporter. . .; TAS.  
 CC GO; GO:0006810; P:transport; TAS.  
 CC InterPro; IPR001898; Na/sul\_sympoort.  
 CC Pfam; PF00939; Na\_sulph\_symp; 1.  
 CC TIGRFAMs; TIGR00785; daas; 1.  
 CC PROSITE; PS01271; NA\_SULFATE; 1.  
 KW Sodium transport; Symport; Transmembrane; Transport.  
 FT TRANSMEM 13 33 Potential.  
 FT TRANSMEM 53 73 Potential.  
 FT TRANSMEM 86 106 Potential.  
 FT TRANSMEM 114 134 Potential.  
 FT TRANSMEM 221 241 Potential.  
 FT TRANSMEM 274 294 Potential.  
 FT TRANSMEM 324 344 Potential.  
 FT TRANSMEM 371 391 Potential.  
 FT TRANSMEM 450 470 Potential.  
 FT TRANSMEM 482 502 Potential.  
 FT TRANSMEM 511 531 Potential.  
 FT TRANSMEM 545 565 Potential.  
 SQ SEQUENCE 592 AA; 64410 MW; 41137D6621A0872A CRC64;

Query Match 78.2%; Score 2392; DB 1; Length 592;  
 Best Local Similarity 76.9%; Pred. No. 1e-172;  
 Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;

Qy 1 MATCPALWAYRFLVILVLCPIFLPLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60  
 Db 1 MATCPALWAYRFLVILVLCPIFLPLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60  
 Qy 61 LFPVLVPLMGIMDASEVCIEYFDKTNILFVGLWVAIAVEHNLKRIALQVLLIIGVR 120



OS Didelphis marsupialis virginiana (North American opossum).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.  
 OX NCBI\_TaxID=9267;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=14973148; DOI=10.1152/ajpcell.00061.2003;  
 RA Aruga S., Pajor A.M., Nakamura K., Liu L., Moe O.W., Preisig P.A.,  
 RA Alpern R.J.,  
 RT "ORP cells express the Na-dicarboxylate cotransporter NaDC-1";  
 RL Am. J. Physiol. Cell Physiol. 287:C64-C72 (2004).  
 DR EMBL; AY186579; AAO27449.1; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006814; P:sodium ion transport; IEA.  
 DR InterPro; IPR001898; Na/sul\_symport.  
 DR Pfam; PF00939; Na sulph sym; 1.  
 DR PROSITE; PS01271; NA\_SULFATE; 1.  
 DR PROSITE; PS01271; NA\_SULFATE; 1.  
 SQ SEQUENCE 605 AA; 67327 MW; 79FAA1AEB6421BE7 CRC64;  
 Query Match 71.2%; Score 2178; DB 2; Length 605;  
 Best Local Similarity 70.0%; Pred. No. 1.7e-156;  
 Matches 421; Conservative 67; Mismatches 93; Indels 20; Gaps 5;  
 QY 1 MATCPALWAYRFLYLVLCLPIFLPLIVQTEKAYCAYSIILMALLWCTEALPLAVTA 60  
 DB 1 MVNLWRLWAYRKYLSILPILLPLPLVPTKEAKCAYSIILMALLWCTEALPLAVTA 60  
 QY 61 LPFVLPLMGIMDASEVCTEYFKDNLFLVGGMLVAIVEHNLHKKRIALQVLLIIGVR 120  
 DB 61 FLPIIFFPMGIMDASEVSEIYKDTNLFTGGLVAIAVEHNLHKKRIALRVLIIIGVR 120  
 QY 121 PALLLGLMVLATFLSWISNTATTMMVPIGHAVLEQLQ--GSKDVEGNNNPTELQ 178  
 DB 121 PAFILGFMVVTAFLSWISNTATTMMVPIAHAVLSQLHKGPEEKDTGHWNISFELQ 180  
 QY 179 EE--CPQKVTKL----DNGOPVSAP-----SEPRTKQTBHRRFSQGLSLCIC 221  
 DB 181 EPHGNPKPEPSLREKENSVPPTSMPEYKEKEBEEKENKEKEHFKLSQGMSLCVC 240  
 QY 222 YSASIGGIATLTGTPNLVLQGVNSLFPONGNVNPFASWFGFAFPTMIILLLLAWLWQ 281  
 DB 241 YASIGGIATLTGTPNLVLQGVNSLFPKNPSVNFASFAGFAFPTMVLILLLSMIWLQ 300  
 QY 282 VFLGVNFRNFGGEGERKQAAFOVIKTYVYLLGLGPMSPFAKTVTVLVLLVVLWFTFR 341  
 DB 301 ILFLGFNFRNFGGRKSQSKERAAQYQVIOTEHKKLGPMSPFAETAVTFLFLLVVLWFTFR 360  
 QY 342 EPGFFPGWGDVTFANEXGQSMASDGTVAIFSLVMFTIPSKIPLGMQDPKPKGLKAPPA 401  
 DB 361 EPGFFPGWGNVAFSNEGSMVSDGTVAIFSLIFVPSLVPFSQO--QGSQKFKAPPA 419  
 QY 402 ILTWKTVNDKPMNIVILLGGFALAKGSGSLSEWLGDKLPLOHIPPSPATAVILCLL 461  
 DB 420 LLDWNTVNMKPMNIVILLGGFALAKGSEVSGLSMWLGNKLPLOHIPPSPATAVILCLL 479  
 QY 462 IAIPECTSNVATTTFLPILASNAQAICLHPVLMVLPCTLAASLAEMLVATPPNAIVF 521  
 DB 480 VATTECANVATTTFLPILASNAQAICLNPVLMVLPCTLSASLAEMLVATPPNAIVF 539  
 QY 522 SFGGLKYSMDARAGFLNIIGVLAITLSINSWSIPIFKLDTFFPSWAHSNTSQCLNPSNS 581  
 DB 540 SYGOLKVIDMAKTGFLNIIGVLTITLAINTWSYPIFQLDQFPWQAQINSTSQVNGQGN 599  
 QY 582 T 582  
 DB 600 T 600  
 RESULT 7  
 OS 057661  
 ID 057661  
 AC 057661;  
 PRELIMINARY; PRT; 622 AA.

DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Intestinal sodium/lithium-dependent dicarboxylate transporter  
 DE (NA(+)/dicarboxylate cotransporter).  
 DE Name=NADC-2;  
 GN Xenopus laevis (African clawed frog).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Intestine;  
 RX MEDLINE=97423461; PubMed=9277403;  
 RA Bai L., Pajor A.M.;  
 RT "Expression cloning of NaDC-2, an intestinal Na(+)- or Li(+)-dependent  
 RT dicarboxylate transporter";  
 RL Am. J. Physiol. 273:G267-G274 (1997).  
 CC -I- FUNCTION: TRANSPORTS DI- AND TRICARBOXYLATES, INCLUDING SUCCINATE,  
 CC CITRATE, AND GLUTARATE, BUT EXCLUDES THE MONOCARBOXYLATE LACTATE.  
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -I- TISSUE SPECIFICITY: Expressed only in the intestine.  
 DR EMBL; U87318; AAB97879.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0015293; F:symporter activity; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006814; P:sodium ion transport; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR001898; Na/sul\_symport.  
 DR Pfam; PF00939; Na sulph sym; 1.  
 DR PROSITE; PS01271; NA\_SULFATE; FALSE NEG.  
 KW Sodium transport; Symport; Transmembrane; Transport.  
 FT TRANSMEM 13 33 POTENTIAL.  
 FT TRANSMEM 53 73 POTENTIAL.  
 FT TRANSMEM 114 134 POTENTIAL.  
 FT TRANSMEM 259 279 POTENTIAL.  
 FT TRANSMEM 299 319 POTENTIAL.  
 FT TRANSMEM 366 386 POTENTIAL.  
 FT TRANSMEM 411 431 POTENTIAL.  
 FT TRANSMEM 456 476 POTENTIAL.  
 FT TRANSMEM 494 514 POTENTIAL.  
 FT TRANSMEM 542 562 POTENTIAL.  
 FT TRANSMEM 584 604 POTENTIAL.  
 FT DOMAIN 123 126 POLY-L-LEU.  
 FT CARBOHYD 617 637 N-LINKED (GLCNAC... (POTENTIAL).  
 SQ SEQUENCE 622 AA; 68550 MW; D25C06BB806B0098 CRC64;  
 Query Match 64.5%; Score 1971.5; DB 2; Length 622;  
 Best Local Similarity 63.2%; Pred. No. 8e-141;  
 Matches 386; Conservative 87; Mismatches 85; Indels 53; Gaps 10;  
 QY 10 AYRFYLVLCPIFLPLPLIVQTEKAYCAYSIILMALLWCTEALPLAVTALPVLFP 69  
 DB 10 ANRNFYIIFLVPLPLPLVPTKEASGFGVIIVNALFWCTEALPLAVTALPVLFP 69  
 QY 70 MGMDASEVCIEFKDNTLFLVGGMLVAIAVEHNLHKKRIALQVLLIIGVRPALLLGLPM 129  
 DB 70 MGMDASTAVCSQYKDKTNMLFGLGLVAISVSKNLHKKRIALRVLIIIGVKPALLLGLPM 129  
 QY 130 LVTAFLSMWISNTATTMMVPIGHAVLEQLQSK-----XDEGGNN----- 171  
 DB 130 VVTAFLSMWISNTATTMMVPIAQAVMQGLHSSEGVDERVEGNSNTQKNVGMENDMYE 189  
 QY 172 -----NPTFELQEC-----POKEVTKLNGQPVSAP--SEPRTKT 206  
 DB 190 SYMPSGKMAIAIDNTYATENEGETQKSTKDPSPSKQ--EKQSIGIVTEPEDEKQTEK 248  
 QY 207 Q--EHRFSGLSLCLCYASIGGIATLTGTPNLVLQGVNSLFPONGNVNPFASWFG 264  
 DB 249 QEKHLKICKMSLVCVCYASIGGIATLTGTPNLVNMKQMDLFPENNININFAFWFG 308  
 QY 265 APFTMIILLLLAWLMLQVLFGLVNFNFRNFGFGEES--ERKQAAFOVIKTYVLLGLPM 322



RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2];  
RP SEQUENCE FROM N.A.  
RC TISSUE=Whole;  
RA Strausberg R.;  
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC063947; AAH63947.1; --  
DR ZFIN; ZDB-GENE-040426-2389; zgc:77607.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006814; P:sodium ion transport; IEA.  
DR InterPro; IPR001898; Na\_sulph\_symp; 1.  
DR Pfam; PF00939; Na\_sulph\_symp; 1.  
SQ SEQUENCE 613 AA; 67653 MW; 8A4B42760D510C6B CRC64;  
  
Query Match 61.1%; Score 1867; DB 2; Length 613;  
Best Local Similarity 59.4%; Pred. No. 6.5e-133;  
Matches 354; Conservative 101; Mismatches 103; Indels 38; Gaps 9;  
  
QY 8 LWAYRFLVILCLPILPLPLIVQTKAYCAYSIIMALLMCTEALPLAVTALPVLVLF 67  
DB 13 LLLHRNYILCLVPLILPLPLVPTPEARCGFAIILMALFWCTECPALITALLPVVLF 72  
  
QY 68 PLMGIMDASEVCIEYKDTNLLFVGLMVAIAVEHNLHKLRIALQVLLIIGVRLPALLIG 127  
DB 73 PMWGIMSEGVQVQLKDTNMLFVGLLVAVAVEHNLHKLRIALSVLLVGVRLPALLMIG 132  
  
QY 128 FMLVTAFLSMWISNTATTAMVPVPIGHVLEQLQGS-----KKDVEGNNNPTFELQE--- 179  
DB 133 FMIVTAFLSMWISNTATTAMVLPISQAVLEQLSATEADSEKELREGVQNOAFELTEVNI 192  
  
QY 180 ECP-----QKEVTKLDNGQVPSAPSPRTQKQE--HHRFSQGLSLCICYSASIGGIAT 231  
DB 193 KQPLDNTQGEKPNNDVVEGINALSE-RRRKAREAKYLRFGMSLSVCYSASIGGTAT 251  
  
QY 232 LTGTTPNLVLQGVNSLFPONGNVNFASFNGFGAFPTMIILLALLMVLQVLFVGNFRK 291  
DB 252 LTGTTPNLILKGQWDEIFPDNDVINNFASFNGFGAFPPNVLMLVSLWNLQCYLGFNFQK 311  
  
QY 292 NFGFGEGERKQAAFOVIKTYQVRLGLGPMGFAEKTVTVLVFLVLMFTREPFGFGWGD 351  
DB 312 SFGCGTKNEGDKO-AYKVMKNEYKKGPMGFAEGAVLVIFLVILWFTREPFGFGWAT 370  
  
QY 352 TVFANEKGQSMASDGTVAIFSLVMFIIIPSKIPGLMQ-----DPKPK 393  
DB 371 ELF-KNKGQ-YVTDGTVAIFPMSTLFFVIPSVDPLCSIKYEODEAVGEQDQVKEKK 428  
  
QY 394 GLKAPPAILTAKVNDKMPNIVILGGGFALAKGSEQSGLEWLDKLTPLQHI;PPSA 453  
DB 429 RKLKGTTLNLKVVHMRPNVILVGGGFALASGEESGLSVWLQOSLPSLPSPPA 488  
  
QY 454 TAVILCLLIAIFTECTSNVATTLFLPILASMAQAICLHPLYVLMFPTLAAISLAFMLPVA 513  
DB 489 ISIILCLLVGTFTECSNTATTLFLPILASMATTIGLHPLYVLMFPTLISASLAFMLPVA 548  
  
QY 514 TPNNAIVFSGGLKVSDMARAGFLIATIGVLAITLSINWSIPIFKLDTFPPSAHS 569  
DB 549 TPNNATAFSYGNLKVLDMAKAGFLIILIGLINCGLINTGWMAFMFKLDTFPPSWANN 604  
  
RESULT 10  
Q803K7  
ID Q803K7 PRELIMINARY; PRT; 613 AA.

AC Q803K7;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Similar to solute carrier family 13, member 2.  
GN ORFNames=zgc:55601;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1];  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB; TISSUE=Whole body;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2];  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB; TISSUE=Whole body;  
RA Strausberg R.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC044437; AAH44437.1; --  
DR ZFIN; ZDB-GENE-040426-2803; zgc:55601.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006814; P:sodium ion transport; IEA.  
DR InterPro; IPR001898; Na\_sulph\_symp; 1.  
DR Pfam; PF00939; Na\_sulph\_symp; 1.  
SQ SEQUENCE 613 AA; 67594 MW; F6B73FE343F288C8 CRC64;  
  
Query Match 60.9%; Score 1862; DB 2; Length 613;  
Best Local Similarity 59.4%; Pred. No. 1.6e-132;  
Matches 354; Conservative 100; Mismatches 104; Indels 38; Gaps 9;  
  
QY 8 LWAYRFLVILCLPILPLPLIVQTKAYCAYSIIMALLMCTEALPLAVTALPVLVLF 67  
DB 13 LLLHRNYILCLVPLILPLPLVPTPEARCGFAIILMALFWCTECPALITALLPVVLF 72  
  
QY 68 PLMGIMDASEVCIEYKDTNLLFVGLMVAIAVEHNLHKLRIALQVLLIIGVRLPALLIG 127  
DB 73 PMWGIMSEGVQVQLKDTNMLFVGLLVAVAVEHNLHKLRIALSVLLVGVRLPALLMIG 132  
  
QY 128 FMLVTAFLSMWISNTATTAMVPVPIGHVLEQLQGS-----KKDVEGNNNPTFELQE--- 179  
DB 133 FMIVTAFLSMWISNTATTAMVLPISQAVLEQLSATEADSEKELREGVQNOAFELTEVNI 192  
  
QY 180 ECP-----QKEVTKLDNGQVPSAPSPRTQKQE--HHRFSQGLSLCICYSASIGGIAT 231  
DB 193 KQPLDNTQGEKPNNDVVEGINALSE-RRRKAREAKYLRFGMSLSVCYSASIGGTAT 251  
  
QY 232 LTGTTPNLVLQGVNSLFPONGNVNFASFNGFGAFPTMIILLALLMVLQVLFVGNFRK 291  
DB 252 LTGTTPNLILKGQWDEIFPDNDVINNFASFNGFGAFPPNVLMLVSLWNLQCYLGFNFQK 311

```
QY 292 NEFGGEEERKQAAFOVKTQYRLGLPMSFAEKTIVLVLLVLLVLTFTREPGFFGWD 351
Db 312 SFGCGTNGEGDKD-AKYVMKNEYKLLGPMSPFAEGLVFIPLVLLVLTFTREPGFPGMAT 370
QY 352 TVFANEKGQSMASDGTVAIFISLVMEIIPSKIPGLMQ-----DPKKP 393
Db 371 ELF-KNGQ-VYTDGTVAIFMSTLFPVIRVDVFLCSIKYERQDEBAVEEODGVEKEK 428
QY 394 GKLLKAPPAILTKVTKVNDKMPNIVILLGGGFALAGSEGLSEWLDKLTPLQHIPPSA 453
Db 429 RLKLGKPTLLNKKVHMERPMNIVILLGGGFALAGSEGLSVMLGQSLSPKSIPTPPA 488
QY 454 TAVILCLLAIETECTSNVATTTLFPIILASNAQAICLHPLYVMLPCTLLASLAFMLPVA 513
Db 489 ISIIICLLVGTTECSNTATTTLFPIILASNAQAICLHPLYVMLPCTLLASLAFMLPVA 548
QY 514 TPNNAIVSFGGLKQVDMARAGFLNIIGVLAITLSINSWSPIPKLDTFFPSWAHS 569
Db 549 TPNNAIVSFGGLKQVDMARAGFLNIIGVLAITLSINSWSPIPKLDTFFPSWAHS 604

RESULT 11
Q6AZR9 PRELIMINARY; PRT; 586 AA.
AC Q6AZR9;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DE 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Slc13a2-protov protein.
GN Name=slc13a2-protov;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC xenopodinae; Xenopus.
ON NCBI_TaxID=8355;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faney J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.B., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Gerhard D.S.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC077435; AAH77435.1; --
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001998; Na/eul_symp; 1.
DR Pfam; PF00939; Na_sulph_symp; 1.
DR PROSITE; PS01271; NA_SULFATE; 1.
SQ SEQUENCE 586 AA; 64959 MW; 2E3646A1D81F31DB CRC64;

Query Match 57.8%; Score 1768.5; DB 2; Length 586;
Best Local Similarity 56.7%; Pred. No. 1.8e-125;
Matches 334; Conservative 103; Mismatches 127; Indels 25; Gaps 6;

QY 1 MATCPALWAFRYLIIVLCPLIFLLPLPLIVQTKAYCAYSIIIMALLMWTALPLAVTA 60
Db 1 MSLCKMKNLCKRFLILFMTPLLLPLLLPLLIQTKKASCAVYIIMAVVYCTEVIPLAVTA 60
QY 61 LFPVILFPLMGIMDASEVCIYFKDNTNLFVGLGMVAIAVEHWNHKLIALQVLLIGVR 120
Db 61 LLPVLLFPFGLFLESKKVCMQYLDKTNMLFVGLIVAVAVEQWNLHKLIALKVLIVGVR 120
QY 121 PALLLLGFMVLTAFLSMWISNTATTAMVPIGHAVLEOLQSGKKDVE-----GGNNNP 175
Db 121 PALLLLGFMVGTAFLSMWSISNTATTAMVPIQVAVLSQLHTAEEDPSMLESLEGQTNPAL 180
QY 176 ELQEE--CPQKEVTKLDNGQPVSPSEPRTQKTOEHHRFSQGLSLCICYSASIGGIATLT 233
Db 181 ELETNAIFPMQLVQTVSNHGVPEVPEVPEKDERKH--TSKGMMLCVCYAASIGGTATLT 238
QY 234 GTTNVLVLOGVNSLFPONGNVNVPASWFGAFPTMILLALLLWLVLLVPLGVNPKRNF 293
Db 239 GTGNVLVKGQFTQTFPNNGDILNFASWFGAFPNMVMILCLISWFLQPSFTFGNFKKTW 298
QY 294 GFGSEERKQAAFOVKTQYRLGLPMSFAEKTIVLVLLVLLVLTFTREPGFFGWD 353
Db 299 CGGATASEKERAAYSVIREYRKLGPIYSASSVLLFILLVLLVLTFTREPGFFGWD 357
QY 354 FANEKGQSMASDGTVAIFISLVMEIIPSKIPGL-----MODPKKPKGLKAPP 400
Db 358 LFNKDNIEYATDATVAVFVALLFILPATKPKGCCARNSFDLEDPEEQKDTFFSAP- 416
QY 401 AILTWTNDKMPNIVILLGGGFALAGSEGLSEWLDKLTPLQHIPPSATAVILCL 460
Db 417 -LLTKVVVQKMPWSIVLLGGGFALAGSDASGLSHWLGQQTPLSHIPPMAIILSL 475
QY 461 LIAITECTSNVATTTLFPIILASNAQAICLHPLYVMLPCTLLASLAFMLVATPPNAIV 520
Db 476 MIAVTECASNVATATLFLPILASMSRSIEVNPYIMIPCTLSTSFAPMLVATPPNAIV 535
QY 521 FSFGGLKQVDMARAGFLNIIGVLAITLSINSWSPIPKLDTFFPSWAHS 569
Db 536 PSYGLHRSVDMVKTGIVWNIIIGIITISINTWGRPMFSLDTFFQWANS 584

RESULT 12
Q86VY5 PRELIMINARY; PRT; 568 AA.
ID Q86VY5;
AC Q86VY5;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Na+-coupled citrate transporter protein.
GN Name=NACT;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22334959; PubMed=12445824; DOI=10.1016/S0006-291X(02)02669-4;
RA Inoue K., Zhuang L., Ganapathy V.;
RT "Human Na+-coupled citrate transporter: primary structure, genomic
RT organization, and transport function.";
```

RL	Biochem. Biophys. Res. Commun. 299:465-471 (2002).	RL	Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AV151633; AAN86530.1; -	DR	EMBL; AK172785; BAD18766.1; -
DR	Genew; HGNC:23089; SLC13A5.	DR	GO; GO:0016020; C-membrane; IEA.
DR	GO; GO:0016020; C-membrane; IEA.	DR	GO; GO:0005215; F-transporter activity; IEA.
DR	GO; GO:0005215; F-transporter activity; IEA.	DR	GO; GO:0006814; P-sodium ion transport; IEA.
DR	GO; GO:0006814; P-sodium ion transport; IEA.	DR	InterPro; IPR001898; Na/sul_symp; 1.
DR	InterPro; IPR001898; Na/sul_symp; 1.	DR	Pfam; PF00939; Na sulph. sym; 1.
DR	Pfam; PF00939; Na sulph. sym; 1.	DR	PROSITE; PS01271; NA_SULFATE; 1.
DR	PROSITE; PS01271; NA_SULFATE; 1.	SQ	SEQUENCE 568 AA; 63062 MW; B8995E56618DECCB CRC64;
SQ	SEQUENCE 568 AA; 63062 MW; B8995E56618DECCB CRC64;		
Query Match 52.1%; Score 1624; DB 2; Length 568;			
Best Local Similarity 53.1%; Pred. No. 1.6e-114;			
Matches 309; Conservative 108; Mismatches 132; Indels 32; Gaps 6;			
QY	1 MATCWPAALWAYRYFLIVLCPIPLPLIVQTKAYCAYSIILMALLWCTEALPLAVTA 60	QY	1 MATCWPAALWAYRYFLIVLCPIPLPLIVQTKAYCAYSIILMALLWCTEALPLAVTA 60
DB	1 MASALSYSVKFKSFVILFVTPLLLPLVILMPAKFVRCAYVILMAIYMCVEIPLAVTS 60	DB	1 MASALSYSVKFKSFVILFVTPLLLPLVILMPAKFVRCAYVILMAIYMCVEIPLAVTS 60
QY	61 LFPVLPFLMGIMDASVCEYFKOTNIFLVGGLMAVAIVEHNLHKLIALQVLLIIGVR 120	QY	61 LFPVLPFLMGIMDASVCEYFKOTNIFLVGGLMAVAIVEHNLHKLIALQVLLIIGVR 120
DB	61 LMPVLLFPLFQILDQRQVCVQYMKDTNMLFLGGLIVAVAVERNLHKLIALRTLLVWGAK 120	DB	61 LMPVLLFPLFQILDQRQVCVQYMKDTNMLFLGGLIVAVAVERNLHKLIALRTLLVWGAK 120
QY	121 PALLLGFMVLVTAFLSMWISNTATTAMVPIGHAVLEQLQGSKKDVEGGNNNFTFLOEE 180	QY	121 PALLLGFMVLVTAFLSMWISNTATTAMVPIGHAVLEQLQGSKKDVEGGNNNFTFLOEE 180
DB	121 PARMLGFMGVTAFLSMWISNTATTAMVPIVEAILQMEATSAATEAG----LELVOK 175	DB	121 PARMLGFMGVTAFLSMWISNTATTAMVPIVEAILQMEATSAATEAG----LELVOK 175
QY	181 CPQKEVTKLDNGQPVASPSPTQKTOEHRFSQGLSLCICYSASIGGIATLTGTTPLNV 240	QY	181 CPQKEVTKLDNGQPVASPSPTQKTOEHRFSQGLSLCICYSASIGGIATLTGTTPLNV 240
DB	176 GKAKE---LPGSQVIFEGPTLQGOEDQERKRLCKAMTLCICYAASIGGTATLTGTGPNV 232	DB	176 GKAKE---LPGSQVIFEGPTLQGOEDQERKRLCKAMTLCICYAASIGGTATLTGTGPNV 232
QY	241 LOGQVNSLPFGNGVNVFASWFGFAPPTMIILLLAWLWQLVFLGVNFRKNGFGEGEE 300	QY	241 LOGQVNSLPFGNGVNVFASWFGFAPPTMIILLLAWLWQLVFLGVNFRKNGFGEGEE 300
DB	233 LLGQNNELFPDSKDLNVNFAWFAFPNMLVLLFAWMLQFVYMRFNFKSGCGLESK 292	DB	233 LLGQNNELFPDSKDLNVNFAWFAFPNMLVLLFAWMLQFVYMRFNFKSGCGLESK 292
QY	301 ERKQAAFOVKTQYRLGPMSPAEKTVTVLFLVLLVWFTREPGFPFGW-----GDTV 353	QY	301 ERKQAAFOVKTQYRLGPMSPAEKTVTVLFLVLLVWFTREPGFPFGW-----GDTV 353
DB	293 KNEKAALKVLEQYRKGLPLSFAEINVLICFFLLVILWFSRDPGFMGMLTVAWVEGETK 352	DB	293 KNEKAALKVLEQYRKGLPLSFAEINVLICFFLLVILWFSRDPGFMGMLTVAWVEGETK 352
QY	354 FANEKGOSMASDGTVAIFISLVNFIPIPSKIPGL-----MQDPKPKGLKAPPAITWKT 407	QY	354 FANEKGOSMASDGTVAIFISLVNFIPIPSKIPGL-----MQDPKPKGLKAPPAITWKT 407
DB	353 Y-----VSDATVAIFVATLFIIVPSQKPNFRSQTEERKTP---FYPPPLLDWKV 401	DB	353 Y-----VSDATVAIFVATLFIIVPSQKPNFRSQTEERKTP---FYPPPLLDWKV 401
QY	408 VNDKMPNVIILGGGFALAKGSEOSGLSEWLGDKLTPLOHIPPSTAVTILCLIIAIFTE 467	QY	408 VNDKMPNVIILGGGFALAKGSEOSGLSEWLGDKLTPLOHIPPSTAVTILCLIIAIFTE 467
DB	402 TOEKVPWGIVLLGGGFALAKGSEASGLSVWGMKQEPHVAVPPAAILTILSLIVAVFTE 461	DB	402 TOEKVPWGIVLLGGGFALAKGSEASGLSVWGMKQEPHVAVPPAAILTILSLIVAVFTE 461
QY	468 CTSNVATTTILFPLASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSFGGLK 527	QY	468 CTSNVATTTILFPLASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSFGGLK 527
DB	462 CTSNVATTTILFPLFASMSRSIGLNPLYIMLPCTLSASFAPMLPVATPPNAIVFTYGHKL 521	DB	462 CTSNVATTTILFPLFASMSRSIGLNPLYIMLPCTLSASFAPMLPVATPPNAIVFTYGHKL 521
QY	528 VSDMARAGFLNIIGVLAITLSINSWSIPFIKLDTFPSWAH 568	QY	528 VSDMARAGFLNIIGVLAITLSINSWSIPFIKLDTFPSWAH 568
DB	522 VADMVKTGVMNIIGVFCVFLAVNTWGRAIFDLDFPDWNAH 562	DB	522 VADMVKTGVMNIIGVFCVFLAVNTWGRAIFDLDFPDWNAH 562
RESULT 13			
Q6ZMG1	PRELIMINARY; PRT; 568 AA.	Q67BT3	PRELIMINARY; PRT; 572 AA.
AC	Q6ZMG1;	AC	Q67BT3;
DT	05-JUL-2004 (TrEMBLrel. 27, Created)	DT	25-OCT-2004 (TrEMBLrel. 28, Created)
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)	DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE	Hypothetical protein FLJ23946.	DE	Sodium-coupled citrate transporter.
OS	Homo sapiens (Human).	OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=9606;		NCBI_TaxID=10090;	
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RA	Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,	RA	Inoue K., Fei Y.-J., Zhuang L., Gopal B., Miyauchi S., Ganapathy V.;
RA	Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,	RT	"Functional features and genomic organization of mouse NaCT, a sodium-
RA	Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;	RT	coupled transporter for citric acid cycle intermediates."



**THIS PAGE IS BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:14 ; Search time 84.655 Seconds  
(without alignments)  
2741.199 Million cell updates/sec

Title: US-10-017-479a-5

Perfect score: 3094

Sequence: 1 MAALAAAKVWSARRLLVL.....HAANVTALPPALTNTVTQL 600

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003s:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2741	88.6	602	6	ABR40097 Human sod
2	2741	88.6	602	6	ABR40101 Human tra
3	2741	88.6	602	6	ABG75837 Transport
4	2326.5	75.2	520	7	ADM05177 Human pro
5	2231.5	72.1	533	8	ABM84303 Human dia
6	1412.5	45.7	616	5	ABU65064 Human NOV
7	1412.5	45.7	616	5	ABU65063 Human NOV
8	1412.5	45.7	616	7	ADK51048 Human NOV
9	1412.5	45.7	616	8	ADH42447 Novel hum
10	1412.5	45.7	616	8	ADN61777 Human nov
11	1412.5	45.7	616	8	ADN61779 Human nov
12	1411.5	45.6	568	6	ABR57023 Human TCH
13	1411.5	45.6	568	7	AAE38764 Human 696
14	1411.5	45.6	568	7	ADK51052 Human NOV
15	1411.5	45.6	568	8	ADH42441 Novel hum
16	1411.5	45.6	568	8	ADP64793 Human Na+
17	1411.5	45.6	576	8	ADH42443 Novel hum
18	1405.5	45.4	568	7	ADE07994 Novel pro
19	1405.5	45.4	619	8	ADG16995 African c
20	1402.5	45.3	568	6	ABG75835 Transport
21	1400.5	45.3	568	5	ABU65062 Human NOV
22	1400.5	45.3	568	7	ADK51050 Human NOV
23	1400.5	45.3	568	8	ADH42445 Novel hum
24	1400.5	45.3	568	8	ADN61775 Human nov
25	1399	45.2	581	8	ADP64799 Zebrafish

99 737 23.8 510 4 AAG90000 C glutami  
100 699.5 22.6 432 6 ABM71452 Staphyloc

ALIGNMENTS

RESULT 1  
ABB82952  
ID ABB82952 standard; protein; 602 AA.  
XX  
AC ABB82952;  
XX  
DT 14-APR-2003 (first entry)  
XX  
DE Human SLC13A related protein (GenBank Identifier No. GI#13653602).  
XX  
KW SLC13A; p53; sodium-sulfate cotransporter 2; cytostatic; cancer;  
KW transmembrane protein; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200298468-A1.  
XX  
PD 12-DEC-2002.  
XX  
PF 03-JUN-2002; 2002WO-US017460.  
XX  
PR 05-JUN-2001; 2001US-0296076P.  
PR 10-OCT-2001; 2001US-0328605P.  
PR 15-FEB-2002; 2002US-0357253P.  
PR 01-MAR-2002; 2002US-0361196P.  
XX  
PA (EXEL-) EXELIXIS INC.

XX  
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;  
PI Lioubin MN;  
XX  
DR WPI; 2003-167297/16.  
XX  
PT Identifying candidate p53 pathway modulating agent for diagnosing or  
PT treating cancer comprises detecting test agent-biased activity of an  
PT assay system comprising purified Sodium sulfate cotransporter 2 (SLC13A)  
PT polypeptide.  
XX  
PS Claim 13; Page 64-67; 69pp; English.

XX  
CC The invention relates to identifying a candidate p53 pathway modulating  
CC agent. The method involves assaying purified Sodium-sulfate cotransporter  
CC 2 (SLC13A) polypeptide or nucleic acid or a functionally active fragment  
CC or derivative with a test agent. The methods are useful for identifying a  
CC candidate p53 pathway modulating agent, modulating a p53 pathway of a  
CC cell, or a mammalian cell and for diagnosing a disease in a patient. The  
CC disease is breast, colon, lung or ovary cancer having greater than 25%  
CC expression level. The method is useful for manufacturing a medicament for  
CC diagnosing or treating breast, colon, lung or ovary cancer. Sequences  
CC ABB82949-952 represent polypeptide sequences related to the human SLC13A  
CC protein

XX  
SQ Sequence 602 AA;  
Query Match 88.6%; Score 2741; DB 6; Length 602;  
Best Local Similarity 87.4%; Pred. No. 1.1e-278;  
Matches 526; Conservative 41; Mismatches 33; Indels 2; Gaps 2;

Qy 1 MAALAAKAKVWSARRLLVLLVPLALLPILFALPPKRGCLYVILLMAVYTCALPLS 60  
Db 1 MAALAAKAKVWSARRLLVLLVPLALLPILFALPPKRGCLYVILLMAVYTCALPLS 60  
Qy 61 VTALLPILPFPMGILPSSKVCPOYFLDTNPLFSLGLIMASAEERNLHRIKALVLMV 120  
Db 61 VTALLPILPFPMGILPSSKVCPOYFLDTNPLFSLGLIMASAEERNLHRIKALVLMV 120

Qy 121 GVQPARLILGMVTTSTFLSMMLNSTASTAMMLPIASAILKSLFGQDTRKDIPLREGDST 180  
Db 121 GVQPARLILGMVTTSTFLSMMLNSTASTAMMLPIANAILKSLFGQKEVRKDPQSSEENT 180  
Qy 181 AAVRGNGLRTVPTQMQLASSEG-GHAEDVEAPLELDDDS-KEEHRRNIWKGFLISIPY 238  
Db 181 AAVRGNGLRTVPTQMQLASTEAKDHPGETEVEPLDLPADSRKEDSYRRNIWKGFLISIPY 240  
Qy 239 SASIGGTATLTGTAPNLILLGOLKSFPPQCDVNVFGSWFIFAPPLMLLFLLVGLWLWISFL 298  
Db 241 SASIGGTATLTGTAPNLILLGOLKSFPPQCDVNVFGSWFIFAPPLMLLFLLVGLWLWISFL 300  
Qy 299 YGMSWRGWRKKNKSLQDAEDKAKAVIQEESQNLGPIKFAEQAVFIFLCLFAILLFSRD 358  
Db 301 YGGLSFRGWRKKNKSEIRTNADBARAVIREEQNLGPIKFAEQAVFIFLCLFAILLFTRD 360  
Qy 359 PKFIPGWASLFPAGFVSDAVTGVAIVTILFFPPSQKPSLKWDFDKAPNSETTEPLLSWKK 418  
Db 361 PKFIPGWASLFPNGFLSDAVTGVAIVTILFFPPSQKPSLKWDFDKAPNTEPLELLTWKK 420  
Qy 419 AQETVPMNIIILLGGGFAMKCESGLSANTGGOLHPLHLEHVPPLAVLLITVIAFFTE 478  
Db 421 AQETVPMNIIILLGGGFAMKCESGLSANTGGOLHPLHLEHVPPLAVLLITVIAFFTE 480  
Qy 479 FASNTATIIIFLPVLAELAIRLHVHPLVLMIPGTVSCSYAFMLPVSTPPNSIAFSTGHLL 538  
Db 481 FASNTATIIIFLPVLAELAIRLHVHPLVLMIPGTVSCSYAFMLPVSTPPNSIAFASGHLL 540  
Qy 539 VKDMVRTGLLMLMGVLLLSLAMNTWAQAI FOLGTFPPDWANTHAANVTALPPALNTNTVQ 598  
Db 541 VKDMVRTGLLMLMGVLLLSLAMNTWAQITFOLGTFPPDWADMYSVNVVTALPPTLANDTFR 600  
Qy 599 TL 600  
Db 601 TL 602  
RESULT 2  
ABR40101  
ID ABR40101 standard; protein; 602 AA.  
XX  
AC ABR40101;  
XX  
DT 24-JUL-2003 (first entry)  
XX  
DE Human renal sodium dicarboxylate co-transporter NADC3.  
XX  
KW Anti-diabetic; anorectic; sodium dicarboxylate cotransporter; human;  
KW sodium tricarboxylate cotransporter; HepNadC; hepatocyte; HepG2;  
KW diabetes; obesity; lipid metabolism; aging; NADC3.  
XX  
OS Homo sapiens.  
XX  
PN WO2003029465-A1.  
XX  
PD 10-APR-2003.  
XX  
PF 27-SEP-2002; 2002WO-JP010038.  
XX  
PR 28-SEP-2001; 2001JP-00299433.  
PR 28-AUG-2002; 2002JP-00249016.  
XX  
PA (SAKA ) OTSUKA PHARM CO LTD.  
XX  
PI Kanemoto N, Omori Y, Sugano S, Obuchi Y;  
XX  
DR WPI; 2003-354728/33.  
XX  
PT Sodium di- or tricarboxylate cotransporter gene (HepNadC) for treatment  
PT and prevention of diabetes, obesity, for improving lipid metabolism and  
PT aging.  
XX  
PS Example 1; Fig 1; 75pp; Japanese.



CC cardiomyopathy, hypertension, asthma, AIDS, allergies, anaemia.  
CC atherosclerosis, atopic dermatitis, diabetes mellitus, osteoarthritis,  
CC osteoporosis, rheumatoid arthritis, psoriasis, infections, trauma,  
CC hepatitis and cancers, including leukemia and lymphoma. The sequences  
CC presented in ABG75819-ABG75844 are the TRICH proteins of the invention  
XX  
SQ Sequence 602 AA;  
  
Query Match 88.6%; Score 2741; DB 6; Length 602;  
Best Local Similarity 87.4%; Pred. No. 1.1e-278;  
Matches 526; Conservative 41; Mismatches 33; Indels .2; Gaps 2;  
  
QY 1 MAALAAKVKWGSARRLLVLLVPLALLPILPALPKGRCGLVILMAVYVWTEALPLS 60  
DB 1 MAALAAKVKWGSARRLLVLLVPLALLPILPALPKGRCGLVILMAVYVWTEALPLS 60  
  
QY 61 VTALLPIILPPFNGILPSSKVCQYFDTNLFSLGLIMASAEERNLHRRALKVLMVLV 120  
DB 61 VTALLPIILPPFNGILPSSKVCQYFDTNLFSLGLIMASAEERNLHRRALKVLMVLV 120  
  
QY 121 GVOPARLILGMVTTTFLSNWLSNTASTAMMLPIASAILKSLFGQDTRKDLREGEDST 180  
DB 121 GVOPARLILGMVTTTFLSNWLSNTASTAMMLPIANAILKSLFGQEVKRDPSQSEENT 180  
  
QY 181 AAVRNGRLTVPTMOPFLASSEG-GHAEDVEAPLELPDDDS-KEEEHRRNIWKGLISIPY 238  
DB 181 AAVRNGRLTVPTMOPFLASSEG-GHAEDVEAPLELPDDDS-KEEEHRRNIWKGLISIPY 240  
  
QY 239 SASIGGTATITGTAPNLILGLQKSPFPQCDVNVNFGSWFIFAPPLMLLFLVGLMWISFL 298  
DB 241 SASIGGTATITGTAPNLILGLQKSPFPQCDVNVNFGSWFIFAPPLMLLFLVGLMWISFL 300  
  
QY 299 YGMSWRGWRKNSKLODVADKAKAVIQEFQNLGPIKFAEQAVRILFCLFAILLFSD 358  
DB 301 YGGLSFRGWRKNSKSEIRTNADRARAVIREEYQNLGPIKFAEQAVRILFCLFAILLFSD 360  
  
QY 359 PKTIPGWASLPAGFVSDAVTGVAIVTILFFPFSQKPSLKWDFKAPNSETBPLLSWKK 418  
DB 361 PKTIPGWASLPAGFVSDAVTGVAIVTILFFPFSQKPSLKWDFKAPNSETBPLLSWKK 420  
  
QY 419 AQETVPWNIIILLGGGFAMAKGCEESGLSAWIGQLHPLSHVPLLAIVLITVVIAPFTE 478  
DB 421 AQETVPWNIIILLGGGFAMAKGCEESGLSAWIGQLHPLSHVPLLAIVLITVVIAPFTE 480  
  
QY 479 FASNTATIIIFLPVLAELAIRLHVHPLYLMIPTGVSCSYAFMLPVSTPPNSIAFSTGHLL 538  
DB 481 FASNTATIIIFLPVLAELAIRLHVHPLYLMIPTGVSCSYAFMLPVSTPPNSIAFSTGHLL 540  
  
QY 539 VKDMVRTGLLMLNMGVLLLSLAMNTWAQAI FOLGTFDPDWMNTAANVTALPPALTNNVTQ 598  
DB 541 VKDMVRTGLLMLNMGVLLLSLAMNTWAQTI FOLGTFDPDWMNTAANVTALPPALTNDTFR 600  
  
QY 599 TL 600  
DB 601 TL 602

RESULT 4  
ADM05177  
ID ADM05177 standard; protein; 520 AA.  
XX  
AC ADM05177;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Human protein of the invention SEQ ID NO:3862.  
XX human; gene therapy; diagnostic marker; pharmaceutical.  
XX Homo sapiens.  
XX  
XX EP1347046-A1.  
XX

PD 24-SEP-2003.  
XX  
XX 12-APR-2002; 2002EP-00008400.  
XX  
XX 22-MAR-2002; 2002JP-00137785.  
XX  
XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
XX Isozaki T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;  
XX  
XX WPI: 2003-723558/69.  
XX N-PSDB; ADM02734.  
XX  
XX New polynucleotides and polypeptides are useful in gene therapy, for  
PT developing a diagnostic marker or medicines for regulating their  
PT expression and activity, or as a target of gene therapy.  
XX  
XX Claim 1; SEQ ID NO 3862; 305pp; English.  
XX  
XX The invention relates to a novel human polynucleotide and the encoded  
CC polypeptide. A polynucleotide of the invention may have a use in gene  
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
CC as a primer for synthesizing the polynucleotide or as a probe for  
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
CC useful in gene therapy, for developing a diagnostic marker or medicines  
CC for regulating their expression and activity, or as a target of gene  
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
CC are useful as pharmaceutical agents. The present sequence represents a  
XX protein sequence of the invention.  
XX  
SQ Sequence 520 AA;  
  
Query Match 75.2%; Score 2326.5; DB 7; Length 520;  
Best Local Similarity 81.4%; Pred. No. 3.7e-235;  
Matches 452; Conservative 36; Mismatches 30; Indels 37; Gaps 3;  
  
QY 48 MAVYWCTEALPLSVTALLPIILPPFMGILPSSKVCQYFDTNLFSLGLIMASAEERN 107  
DB 1 MAVYWCTEALPLSVTALLPIILPPFMGILPSSKVCQYFDTNLFSLGLIMASAEERN 60  
  
QY 108 LHRRIALKVLMVGVOPARLILGMVTTTFLSNWLSNTASTAMMLPIASAILKSLFGQD 167  
DB 61 LHRRIALKVLMVGVOPARLILGMVTTTFLSNWLSNTASTAMMLPIANAILKSLFGQKE 120  
  
QY 168 TRKDLREGEDSTAAVRNGRLTVPTMOPFLASSEG-GHAEDVEAPLELPDDDS-KEEEHR 225  
DB 121 VRKDPQSEENTAAVRNGRLTVPTMOPFLASSEG-GHAEDVEAPLELPDDDS-KEEEHR 180  
  
QY 226 RNIWKGLISIPYSASIGGTATITGTAPNLILGLQKSPFPQCDVNVNFGSWFIFAPPLML 285  
DB 181 RNIWKGLISIPYSASIGGTATITGTAPNLILGLQKSPFPQCDVNVNFGSWFIFAPPLML 240  
  
QY 286 LFLVGLMWISFLYGGMSWRGWRKNSKLODVADKAKAVIQEFQNLGPIKFAEQAVFI 345  
DB 241 LFLVGLMWISFLYGGMSWRGWRKNSKSEIRTNADRARAVIREEYQNLGPIK 292  
  
QY 346 LFCLFAILLFSDRDKPIPGWASLPAGFVSDAVTGVAIVTILFFPFSQKPSLKWDFDKA 405  
DB 293 -----FLSDAVTGVAIVTILFFPFSQKPSLKWDFDKA 325  
  
QY 406 PNSETBPLLSWKKAEQETVPWNIIILLGGGFAMAKGCEESGLSAWIGQLHPLSHVPLLA 465  
DB 326 PNTETEPILLTWKKAQETVPWNIIILLGGGFAMAKGCEESGLSAWIGQLHPLSHVPLLA 385  
  
QY 466 VLLITVVIAPFTEFASNTATIIIFLPVLAELAIRLHVHPLYLMIPTGVSCSYAFMLPVST 525  
DB 386 VLLITVVIAPFTEFASNTATIIIFLPVLAELAIRLHVHPLYLMIPTGVSCSYAFMLPVST 445  
  
QY 526 PPNISAFSTGHLLAVKDMVRTGLLMLNMGVLLLSLAMNTWAQAI FOLGTFDPDWMNTAANVT 585  
DB 446 PPNISAFSGHLLAVKDMVRTGLLMLNMGVLLLSLAMNTWAQTI FOLGTFDPDWMNTAANVT 505

QY 586 TALPPALNTNTVQTL 600  
 DB 506 TALPPTLANDTPTTL 520

RESULT 5  
 ABM84303  
 ID ABM84303 standard; protein; 533 AA.  
 AC  
 XX  
 XX  
 DT 18-NOV-2004 (first entry)  
 DE Human diagnostic and therapeutic pprotein SEQ ID NO:4552.  
 DE Human diagnostic and therapeutic polynucleotide; dithp.  
 KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
 OS Homo sapiens.  
 XX  
 XX WO2004023973-A2.  
 PN 25-MAR-2004.  
 PD 12-SEP-2003; 2003WO-US028227.  
 PF 12-SEP-2002; 2002US-0410259P.  
 PR 12-SEP-2002; 2002US-0410260P.  
 XX (INCY-) INCYTE CORP.  
 PA Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
 PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Blder LV;  
 PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;  
 PI Stevens KA, Blanchard JB, Panzer SR, Wang X, Au AP, Gerstin EH;  
 PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;  
 PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;  
 PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
 PI Patuary S, Shi X, Suarez CJ;  
 XX WPI: 2004-329368/30.  
 DR N-PSDB; ACN42955.  
 XX  
 XX New diagnostic and therapeutic polynucleotides and polypeptides, useful  
 PT in diagnosing a condition, disease or disorder associated with human  
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
 PT in gene mapping.  
 XX  
 PS Claim 27; Page: 190pp; English.

The invention relates to novel diagnostic and therapeutic polynucleotides  
 CC selected from one of the 2722 sequences defined in the specification. A  
 CC polynucleotide of the invention may have a use in gene therapy. The human  
 CC polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be  
 CC used to diagnose a particular condition, disease or disorder associated  
 CC with human molecules, e.g. cell proliferative disorders,  
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
 CC disorder, neurological disorders, gastrointestinal disorders, or  
 CC infections caused by virus, bacteria, fungi or parasite. The dithp  
 CC molecules may also be used in genetic mapping, in identifying individuals  
 CC from minute biological samples, in detecting single nucleotide  
 CC polymorphisms, as molecular weight markers, and for somatic or germline  
 CC gene therapy. The present sequence represents a dithp protein of the  
 CC invention. Note: The sequence data for this patent is not represented in  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
 XX

Sequence 533 AA;  
 Query Match 72.1%; Score 2231.5; DB 8; Length 533;  
 Best Local Similarity 73.8%; Pred. No. 3.8e-225;  
 Matches 445; Conservative 39; Mismatches 46; Indels 73; Gaps 5;

QY 1 MAALAAAKKWSARRLLVLLVPLALLPILFALPPKGGRCGLYVILLMAVYCTEALPLS 60  
 DB 1 MAALAAAKKWSARRLLVLLVPLALLPILFALPPKGGRCGLYVILLMAVYCTEALPLS 59  
 QY 61 VTALLPIILFPFMGILPSSKVCPOVFLDTNLFSLGLIMASAEIRNLHRRHRTALVLMV 120  
 DB 60 VTALLPIILFPFMGILPSSKVCPOVFLDTNLFSLGLIMASAEIRNLHRRHRTALVLMV 119  
 QY 121 GVQPARLILGMVTTTSFLSMWLSNTASTAMMLPIASAILKSLFGQDRTRKDLPRGEDST 180  
 DB 120 GVQPARLILGMVTTTSFLSMWLSNTASTAMMLPIANAILKSLFGQKEVRKQDSQSEENT 179  
 QY 181 AAVRGNGLRTV-PTMQFLASSEG-GHAEDVEAPLELPDDS-KEBEHRNINWKGFLISIP 237  
 DB 180 GEAVGETAYTLCPRRCSPXPTEAKDHFGETEVLPLDLPADSRKDEYRRNINWKGFLISIP 239  
 QY 238 YSASIGGTATLTGTAPNLILGOLKSRFPQCDVNVFGSWFIFAFPLMLLFLLVGLMWISF 297  
 DB 240 YSASIGGTATLTGTAPNLILGOLKSRFPQCDVNVFGSWFIFAFPLMLLFLLVGLMWISF 299  
 QY 298 LYGGMSWRGMRKQNSKLQDVAEDKAKAVIQBEFQNLGPIKFAEQAVFILFCLFALLPSR 357  
 DB 300 LYGGLSFRGMRKNSKSEIRTNADRARAVIREYQNLGP----- 337  
 QY 358 DKFIPGWASLFAFGFVSDAVTGVAIVTILFFFPKPSLKWDFDKAPNSETPELLSKW 417  
 DB 338 -----APNTETEPLLTAK 350  
 QY 418 KAQETVPWNIIILGGGFAMAKGCEESGLSAGIQLHPLHVPPLLAIVLITVIAFFT 477  
 DB 351 KAQETVPWNIIILGGGFAMAKGCEESGLSAGIQLHPLHVPPLLAIVLITVIAFFT 410  
 QY 478 EFASNTATIIIFLPVLAELAIRLHVHPLYLMIPGTVCSCYAFMLPVSTPPNSIAFSTGHL 537  
 DB 411 EFASNTATIIIFLPVLAELAIRLHVHPLYLMIPGTVCSCYAFMLPVSTPPNSIAFSTGHL 470  
 QY 538 LVKDMVTRGLMNLGMVLLSLAMNTWAQAIIFQLGTFPDWANTHAANYTALPPALTNTV 597  
 DB 471 LVKDMVTRGLMNLGMVLLSLAMNTWAQAIIFQLGTFPDWANTHAANYTALPPALTNTV 530  
 QY 598 QTL 600  
 DB 531 RTL 533

RESULT 6  
 ABU65064  
 ID ABU65064 standard; protein; 616 AA.  
 AC ABU65064;  
 XX  
 XX 20-MAY-2003 (first entry)  
 DT  
 DE Human NOV14c protein.  
 XX NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;  
 KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;  
 KW human.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200272757-A2.  
 XX  
 PD 19-SEP-2002.  
 XX  
 PF 08-MAR-2002; 2002WO-US006908.  
 XX  
 XX 08-MAR-2001; 2001US-0274101P.  
 PR 08-MAR-2001; 2001US-0274194P.  
 PR 08-MAR-2001; 2001US-0274281P.  
 PR 08-MAR-2001; 2001US-0274322P.  
 PR 09-MAR-2001; 2001US-0274849P.  
 PR 12-MAR-2001; 2001US-0275235P.

PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276776P.  
PR 19-MAR-2001; 2001US-0276994P.  
PR 20-MAR-2001; 2001US-0277329P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 21-MAR-2001; 2001US-0277791P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278999P.  
PR 27-MAR-2001; 2001US-0279036P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0277338P.  
PR 30-MAR-2001; 2001US-0279995P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
PR 02-APR-2001; 2001US-0280822P.  
PR 02-APR-2001; 2001US-0280900P.  
PR 04-APR-2001; 2001US-0281194P.  
PR 13-APR-2001; 2001US-0283675P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288528P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 30-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294999P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
PR 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
PR 18-OCT-2001; 2001US-0330380P.  
PR 31-OCT-2001; 2001US-0335301P.  
PR 14-NOV-2001; 2001US-0332172P.  
PR 14-NOV-2001; 2001US-0332271P.  
PR 14-NOV-2001; 2001US-0332272P.  
PR 14-NOV-2001; 2001US-0333184P.  
PR 14-NOV-2001; 2001US-0333272P.  
PR 21-NOV-2001; 2001US-0332094P.  
PR 03-DEC-2001; 2001US-0337426P.  
PR 03-DEC-2001; 2001US-0338092P.  
PR 04-DEC-2001; 2001US-0337185P.  
PR 03-JAN-2002; 2002US-0345705P.  
PR 07-MAR-2002; 2002US-00092900.  
XX  
XX (CURA-) CURAGEN CORP.  
XX Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;  
XX Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;  
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;  
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;  
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;  
PI Lepley DM, Rieger DK;  
XX  
XX WPI; 2002-723332/78.  
DR N-PSDB; ABX97031.  
XX  
XX NOVX polypeptides and polynucleotides, useful for preventing or treating  
PT a disorder associated with aberrant NOVX expression or activity e.g.,  
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial

PT asthma.  
XX  
XX Claim 1; Page 134; 1103pp; English.  
CC This invention describes novel human NOVX polypeptides which have  
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive  
CC activity. Pharmaceutical compositions comprising the NOVX proteins or  
CC nucleic acid molecules or NOVX antibodies are useful for preventing or  
CC treating a disorder associated with aberrant NOVX expression or activity  
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial  
CC asthma. The products of the invention can be used for gene therapy or in  
CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by  
XX ABX97008-ABX97185  
SQ Sequence 616 AA;  
Query Match 45.7%; Score 1412.5; DB 5; Length 616;  
Best Local Similarity 46.6%; Pred. No. 7.4e-139;  
Matches 292; Conservative 102; Mismatches 171; Indels 61; Gaps 13;  
QY 4 LAALAKYWSARRLLVLLVPLALLPILFAPKPKGRCLYVILLMAVYVWCTEALPLSVTA 63  
DB 1 MASALSYVSKPKSFVILFVTPLLPLLVILMPAKVSCCAYVILMAIYCTEVIPLAVIS 60  
QY 64 LLPILFPFPMGILPSPKVCQYPLDTNLFSLGLIMASAEERNLHRRALKVLMVGVQ 123  
DB 61 LMPVLLFPLFQILDQRQVCQYMKDTNMLFLGLIVAVAVERNLHKRIALRTLLVVGAK 120  
QY 124 PARLLGMMVTTSFLSMWLSNTASTAMMLPIASAILK-----SLPGQDTRKDL 172  
DB 121 PARLLMGFGVLTALLSMWISNTATTAMVPIVEAILQQMEATSAEAGLEGQGTINNL 180  
QY 173 PREGEDSTAARVGNGLRVTPTQMQLASSEGHAEDVEAPLE-----LPDD 218  
DB 181 NALEDTVKAVLGG--KCVALLSTYVYKVKELQINNMTPLKLEQEQDGLGPIRPQD 238  
QY 219 S---KEEHRHNIWGLFSLIPYSASIGTATLTGTAPNLILGLQKSPFQPC-DVYNFG 274  
DB 239 SAQCQEDQERKRLCKAMTLCICVAASIGGTATLTGTGNVVLGQNMELFPDPSKDLVNEA 298  
QY 275 SWFIFAPFLMLFLVGLWLSIFLYGMSW-RGW-----RKNKSLQDVAEDKAKAVIOE 328  
DB 299 SWFAPFAPNNLVMLFPAWLQVYVNFSSFKKSWGCGLESKKNEK-----AALKVLQE 351  
QY 329 EFQNLGPIKFAEQAVFELFCLFAILLFSRDPKFIQWASL-FAPG---FVSDAVTGVAVI 384  
DB 352 EYKLGPLSFAEINVLICFLLVILWFSRDPGPMGMLTVAVWEGETKSVDATVAIFVA 411  
QY 385 TILFFPFSQKPSLKWDFDKAPNSETE-----PLLSWKKAOETVPNNIILGGGFA 436  
DB 412 TLLFIVPSQPKP----FNFRSQTEEGKSPVLIAPPFLDWMKVQEKVPNGIVLLGGGFA 467  
QY 437 MAKCEESGLSAWIGQLHPLHVPPLAVLLITVIAPTFPASNTATIIIFLPVLAEL 496  
DB 468 LAKGSEASGLSVWVGQMEPLHAPPAATLILSLVAVFTECTSNVATTLFELPFSAM 527  
QY 497 AIRLHVHPLYMLIPGTVSCSYAFMLPVSTPPNSIAFSTGHLHVKDMVTRTGLLNLKGVLL 556  
DB 528 SRSIGLNPLYIMLPCTLSASFAPMLPVATPPNNAIVFTYGYHLKVADVMKVTGVIINNIIGVFC 587  
QY 557 LSLAMNTWAQAIPLQGTFFPDWAN-TH 581  
DB 588 VFLAVNTWGRAIFDLDFPDWANVTH 613  
RESULT 7  
ID ABU65063 standard; protein; 616 AA.  
XX  
XX AC ABU65063;  
XX  
XX -DT 20-MAY-2003 (first entry)  
XX



Db 468 LAGSEASGLSVWVGKQMEPLHVPAAITLILSLVAVFTECTSNVATTTLFLPFAQM 527  
QY 497 AIRLHVHPLYLMPGTVCSCSYAFMPLPVSTPPNSIAFSTGHLVYKQWVRTGLANLNGVLL 556  
Db 528 SRSIGLNPPLYLMPGTVCSCSYAFMPLPVSTPPNSIAFSTGHLVYKQWVRTGLANLNGVLL 587  
QY 557 LSLAMNTWAQAIIFOLGTFPPDWMAN-TH 581  
Db 588 VFLAVNTWGRAIFDLDHFPDWMANVTH 613

RESULT 8  
ADK51048  
ID ADK51048 standard; protein; 616 AA.  
XX  
AC ADK51048;  
DT 17-JUN-2004 (first entry)  
XX  
DE Human NOV18A protein sequence SeqID68.  
XX  
KW cytosolic; NOVX-agonist; NOVX-antagonist; vaccine; gene therapy; cancer;  
KW chromosome mapping; human; NOV18A.  
XX  
OS Homo sapiens.  
XX  
PN WO2003083046-A2.  
XX  
PD 09-OCT-2003.  
XX  
PF 01-APR-2003; 2003WO-US010142.  
XX  
PR 02-APR-2002; 2002US-00115479.  
PR 05-APR-2002; 2002US-0370349P.  
PR 08-APR-2002; 2002US-0370969P.  
PR 12-APR-2002; 2002US-0372019P.  
PR 22-APR-2002; 2002US-0374379P.  
PR 30-MAY-2002; 2002US-0384543P.  
PR 03-JUN-2002; 2002US-00160619.  
PR 15-AUG-2002; 2002US-0403748P.  
PR 04-NOV-2002; 2002US-00287226.  
PR 31-MAR-2003; 2003US-00403161.  
XX  
FA (CURA-) CURAGEN CORP.  
XX  
PI Anderson DW, Bento P, Boldog FL, Burgess CE, Casman SJ, Furtak K;  
PI Gorman L, Gould-Rothberg BE, Gunther E, Heyes MP, Li L, Spytek KA;  
PI Stone DJ, Zhong M, Malyankar UM, Edinger SR, Patturajan M;  
PI Rothenberg ME, Smithson G;  
XX  
DR WPI; 2003-812539/76.  
DR N-PSDB; ADK51047.  
XX  
PT New NOVX polypeptide, useful for preparing a composition for treating or  
PT preventing e.g. cancer or for chromosome mapping.  
XX  
PS Claim 1; SEQ ID NO 68; 433pp; English.  
XX  
CC This invention relates to novel isolated polypeptides and the DNA  
CC sequences which encode them. The invention may be useful for the  
CC development of compounds with a cytostatic activity (as NOVX-agonists or  
CC antagonists) or vaccines. In addition, the disclosed sequences may be  
CC useful for gene therapy. The polypeptide is useful for preparing a  
CC composition for treating or preventing a pathological state in a mammal,  
CC for example cancer or for chromosome mapping. The present sequence is  
CC that of a human NOVX protein of the invention.  
XX  
SQ Sequence 616 AA;

Query Match 45.7%; Score 1412.5; DB 7; Length 616;  
Best Local Similarity 46.6%; Pred. No. 7.4e-139;  
Matches 292; Conservative 102; Mismatches 171; Indels 61; Gaps 13;

QY 4 LAALAKVMSARRLLVLLVPLALPILFALPPKEGRCLYVILLMAVYMWCTEALPLSVTA 63  
Db 1 MASALSYSKPKSFVILFVTPALLPLVILMPAKVSCCAYVILMAIYMWCTEVIPLAVTS 60  
QY 64 LPIILPFPFGMILPSKVCQYFLDTNLFSLGLMASAIBERNLHRRALKVILMLVGVO 123  
Db 61 LMPVLLFPLFQILDSRQVCVQYMKDNTMLFLGGLIVAVAVERNLHRRALKVILMLVGAK 120  
QY 124 PARLILGNVVTLSFLSMLSNTASTAMMLPLASATLK-----SLFQORDTRKDL 172  
Db 121 PARLMLGFMGVTTALLSMMISNTATTAMVPIVEALQQMEATSAAEAGLEGCGTTINNL 180  
QY 173 PREGEDSTAARVNGRLRTVPTMQLASSEGHAEDVEAPLE-----LPDD 218  
Db 181 NALEDVTVKAVLGG--KCVAILSTVYKVKELQINNLTFLKLEKQEQDQDLPGRIPQD 238  
QY 219 S---KEEHRNRNIWGLFISIPYSASIGGTATLTGTAPNLIILGQLKSFPOC-DVWNFG 274  
Db 239 SAQCQEDQERKLCRKLCKAMTLCICAAASIGGTATLTGTGNVLLGQNMELFPPSKDLVFA 298  
QY 275 SWFIFAPPLMLFLVGLWISFLYGGMSW-RGW-----RKNSKLQDVAEDKAKAVIOE 328  
Db 299 SWFAFAPNMLVNLFLFAMLWLFQVYMFSSFKSWGCGLESKNEK-----AALKVLQE 351  
QY 329 EFQNLGPIKFAEQAVFIFLCLFALLFSRDPKPIPGWASL-FAPG---FVSDAVTGVAVI 384  
Db 352 EYRKLGLPLSFAEINVLICFLLVILWFSRDPGMPGWLTVAVVEGETKSVSDATVAIFVA 411  
QY 385 TILFPPPSQKSLKWFDFKAPNSETE-----PLLSWKKAQETVPNNIILLLGGGFA 436  
Db 412 TLLFIVPSQPKP-----FNFRSQTEGKSPVLIAPPPLLDKMTQEKVPNGI VLLLLGGGFA 467  
QY 437 MAGCEESGLSAWIGGOLHPLHVPPLAVLITVIAFFTFBASNTATIIIFLPVLAEL 496  
Db 468 LAKSEASGLSVWVGKQMEPLHVPAAITLILSLVAVFTECTSNVATTTLFLPFAQM 527  
QY 497 AIRLHVHPLYLMPGTVCSCSYAFMPLPVSTPPNSIAFSTGHLVYKQWVRTGLANLNGVLL 556  
Db 528 SRSIGLNPPLYLMPGTVCSCSYAFMPLPVSTPPNSIAFSTGHLVYKQWVRTGLANLNGVLL 587  
QY 557 LSLAMNTWAQAIIFOLGTFPPDWMAN-TH 581  
Db 588 VFLAVNTWGRAIFDLDHFPDWMANVTH 613

RESULT 9  
ADH42447  
ID ADH42447 standard; protein; 616 AA.  
XX  
AC ADH42447;  
XX  
DT 25-MAR-2004 (first entry)  
XX  
DE Novel human protein NOV67d.  
XX  
KW cardiovascular; antiarteriosclerotic; hypotensive; cytostatic; anorectic;  
KW antidiabetic; immunosuppressive; anti-HIV; neuroprotective; nootropic;  
KW antiparkinsonian; antiasthmatic; antiinfertility; cardiomyopathy;  
KW atherosclerosis; hypertension; cancer; obesity; diabetes; AIDS;  
KW multiple sclerosis; graft-versus-host disease; Alzheimer's disease;  
KW Parkinson's disease; asthma; fertility disorder; chromosome mapping;  
KW tissue typing; preventive medicine; pharmacogenomic; vaccine.  
XX  
OS Homo sapiens.  
XX  
PN WO2003102159-A2.  
XX  
PD 11-DEC-2003.  
XX  
PF 04-JUN-2003; 2003WO-US017573.  
XX  
PR 04-JUN-2002; 2002US-0385490P.

PR	04-JUN-2002;	2002US-0385615P.	XX	XX	New NOVX polypeptides and nucleic acid molecules useful for preventing or
PR	04-JUN-2002;	2002US-0385755P.	PT	PT	treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PR	05-JUN-2002;	2002US-0386041P.	PT	PT	atherosclerosis or diabetes, in chromosome mapping, tissue typing or in
PR	06-JUN-2002;	2002US-0386355P.	PT	PT	pharmacogenomics.
PR	06-JUN-2002;	2002US-0386357P.	XX	XX	Claim 2; SEQ ID NO 1000; 1503pp; English.
PR	06-JUN-2002;	2002US-0386447P.	XX	XX	The invention relates to 566 new isolated human polypeptides and their
PR	06-JUN-2002;	2002US-0386459P.	CC	CC	encoding genes, sequences to 566 new isolated human polypeptides and their
PR	06-JUN-2002;	2002US-0386465P.	CC	CC	sequences comprising one or more conservative substitutions in these or
PR	06-JUN-2002;	2002US-038664P.	CC	CC	polypeptide, polynucleotide and antibodies against the polypeptides are
PR	07-JUN-2002;	2002US-0386701P.	CC	CC	useful in diagnosing, treating or preventing NOVX-associated disorders,
PR	07-JUN-2002;	2002US-0386796P.	CC	CC	e.g. cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,
PR	07-JUN-2002;	2002US-0386931P.	CC	CC	Alzheimer's disease, Parkinson's disease, asthma, or fertility disorders.
PR	07-JUN-2002;	2002US-0387078P.	CC	CC	The nucleic acids are further used as hybridization probes, in chromosome
PR	07-JUN-2002;	2002US-0387081P.	CC	CC	mapping, tissue typing, preventive medicine, and pharmacogenomics. The
PR	07-JUN-2002;	2002US-0387083P.	CC	CC	polypeptides are also useful as vaccines. This sequence represents an
PR	10-JUN-2002;	2002US-0387429P.	XX	XX	example of the polypeptide of the invention.
PR	10-JUN-2002;	2002US-0387540P.	XX	XX	Sequence 616 AA;
PR	10-JUN-2002;	2002US-0387866P.	QY	QY	Query Match 45.7%; Score 1412.5; DB 8; Length 616;
PR	11-JUN-2002;	2002US-0387606P.	Db	Db	Best Local Similarity 46.6%; Pred. No. 7.4e-139;
PR	11-JUN-2002;	2002US-0387610P.	QY	QY	Matches 292; Conservative 102; Mismatches 171; Indels 61; Gaps 13;
PR	11-JUN-2002;	2002US-0387659P.	Db	Db	4 LAALAKKWSARRLLVLLVPLALLPILFALPPKGRCLYVILLMAVWCTEALPLSVTA 63
PR	11-JUN-2002;	2002US-0387859P.	QY	QY	1 MASALSYSVKFSFVILFVTELLPLVLLPILPAKVSOCAYVILLMAVWCTEIVPLVTS 60
PR	12-JUN-2002;	2002US-0387960P.	Db	Db	64 LLPILFPFMGILPSSKVCPOYFLDTNPLFLSLGSLIMASAIERNLHRRALKVLMLGVQ 123
PR	12-JUN-2002;	2002US-0388096P.	QY	QY	61 LMPVLLPFLQILDQRVCVQYMDTNMLFLGGLIVAVAVERNLHRRALRTLLWVGAK 120
PR	12-JUN-2002;	2002US-0388432P.	Db	Db	124 PARLILGMVTTSLFSLMWSNTASTAMMLPIASAILK-----SLFGQRTKDL 172
PR	12-JUN-2002;	2002US-0388479P.	QY	QY	121 PARMLGFMGVTTALLSMWISNTATTAMVPIVEAILQOMEATSAATEAGLEGGTTINL 180
PR	13-JUN-2002;	2002US-0389123P.	Db	Db	173 PREGEDSTAAVRGNGLRTVPTMOPFLASSEGHAEDVEAPLE-----LPDD 218
PR	14-JUN-2002;	2002US-0389120P.	QY	QY	181 NALSDDTVKVAVLGG--KCVAIISTYVKVKEKQINLMTPLKKLEKQDQGLGPIRQD 238
PR	14-JUN-2002;	2002US-0389146P.	Db	Db	219 S---KEBEHRRNIWKGLFISIPYSAGSGTATLGTAPNLILLGLQKSFPPQC-DVNVFG 274
PR	17-JUN-2002;	2002US-0389742P.	QY	QY	239 SAQCOEDQERKLCCKAMTLCIYASISGGTATLGTGTVLLGMNLPDPSKDLVNA 298
PR	18-JUN-2002;	2002US-0389604P.	Db	Db	275 SWFIFAPPLMLLFLVGLWLTISFLYGGMSW-RGW-----RKNSKLQDVAEDKAKAVIQE 328
PR	19-JUN-2002;	2002US-0390066P.	QY	QY	299 SWFAPAPNMLVLLFAWMLQFVYMFSSFKKSGCGLESKKNEK-----AALKVLOE 351
PR	19-JUN-2002;	2002US-0390144P.	Db	Db	329 EFQNLGPIKFAEOAVFLLFCLFALLPSRDPKFIPIGNASL-FAPG---FVSDAVTGVAV 384
PR	25-JUN-2002;	2002US-0391726P.	QY	QY	352 EYKLGPLSPAEINVLICFFLLVILNFSRDPGFPGMLTVAVVEGETSKVSADATVAIFVA 411
PR	06-AUG-2002;	2002US-0401628P.	Db	Db	385 TILFFPSQPSLKWDFKAPNSETE-----PLLSWKKAOETVPWNITILLGGGFA 436
PR	09-AUG-2002;	2002US-0402268P.	QY	QY	412 TLLFIVPSQPK----FNFRSQTEEGSKPVLIAPPLLDWKVQKVPWGVILLGGGFA 467
PR	12-AUG-2002;	2002US-0402822P.	Db	Db	437 MAKCESSGSLAWTGGQLHPLHVPPLAVLLITVIAFFTEFASNTATIIIFLPVLAE 496
PR	13-AUG-2002;	2002US-0403458P.	QY	QY	468 LAGSEASGLSVWVGKQWELHVPVPAATILLSLLVAVFTECTSNVATTTFLPIFASM 527
PR	15-AUG-2002;	2002US-0403617P.	Db	Db	497 AIRLHVHLYIMIPCTVSCSYAEMLPYSTPNSTAFSTGHLLVKDMVTVTGLLMNLMVLL 556
PR	26-AUG-2002;	2002US-0406182P.	QY	QY	528 SRSIGLNPILMIFCTLSASFAMLPVATPPNAIVFTYGHLLKVDAMVTKVTGIMNIIIGVFC 587
PR	12-SEP-2002;	2002US-0410083P.	Db	Db	557 LSLAMNTWAQAI FOLGTFFPDWAN-TH 581
PR	23-SEP-2002;	2002US-0410505P.	QY	QY	588 VFLAVTWGRAIFDLHDHFPDWNVTH 613
PR	30-SEP-2002;	2002US-0412955P.	XX	XX	
PR	23-OCT-2002;	2002US-0420627P.	XX	XX	
PR	24-OCT-2002;	2002US-0420718P.	XX	XX	
PR	24-OCT-2002;	2002US-0420852P.	XX	XX	
PR	31-OCT-2002;	2002US-0422750P.	XX	XX	
PR	01-NOV-2002;	2002US-0423095P.	XX	XX	
PR	05-NOV-2002;	2002US-0423748P.	XX	XX	
PR	(CURA-) CURAGEN CORP.		XX	XX	
PI	Alsbrook JP, Anderson DW, Baumgartner JC, Berghs C, Boldog FL;		PI	PI	
PI	Burgess CE, Casman SJ, Catterton E, Dhanabal M, Edinger SR;		PI	PI	
PI	Ellerman K, Ettenberg S, Gangolli EA, Gerlach VJ, Gorman L;		PI	PI	
PI	Grosse WM, Gunther E, Guo X, Gusev VY, Hermann JL, Ji W, Kekuda R;		PI	PI	
PI	Khrantsov NV, Larochele WJ, Li L, Liang H, Low K, Macdougall JR;		PI	PI	
PI	MacLachlan T, Malayanar K, McQueeney K, Mezick AJ, Miller CE;		PI	PI	
PI	Millet I, Padigaru M, Patturajan M, Peyman JA, Qian X, Raetelli L;		PI	PI	
PI	Rieger DK, Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G;		PI	PI	
PI	Spytek KA, Stone DJ, Sukumaran S, Szekeres ES, Vernet CAM, Voss EZ;		PI	PI	
PI	Wolenc AR, Zhong M, Zhong H;		PI	PI	
XX	WPI; 2004-053467/05.		XX	XX	
DR	N-PSDB; ADH42446.		DR	DR	

ADN61777  
ID ADN61777 standard; protein; 616 AA.  
XX  
AC ADN61777;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Human novel protein NOV14b.  
XX  
XX Human; NOVX; diabetes; obesity; infectious disease; anorexia;  
KW cancer-associated cachexia; cancer; neurodegenerative disorder;  
KW Alzheimer's disease; Parkinson's disease; immune disorder;  
KW haematopoietic disorder; dyslipidaemia; chronic disease.  
XX  
OS Homo sapiens.  
XX  
PN US2004043382-A1.  
XX  
PD 04-MAR-2004.  
XX  
XX 07-MAR-2002; 2002US-00092900.  
XX  
PR 08-MAR-2001; 2001US-0274191P.  
PR 08-MAR-2001; 2001US-0274194P.  
PR 08-MAR-2001; 2001US-0274281P.  
PR 08-MAR-2001; 2001US-0274322P.  
PR 09-MAR-2001; 2001US-0274849P.  
PR 12-MAR-2001; 2001US-0275235P.  
PR 13-MAR-2001; 2001US-0275578P.  
PR 13-MAR-2001; 2001US-0275579P.  
PR 13-MAR-2001; 2001US-0275601P.  
PR 14-MAR-2001; 2001US-0276000P.  
PR 16-MAR-2001; 2001US-0276776P.  
PR 19-MAR-2001; 2001US-0276994P.  
PR 20-MAR-2001; 2001US-0277239P.  
PR 20-MAR-2001; 2001US-0277321P.  
PR 20-MAR-2001; 2001US-0277327P.  
PR 20-MAR-2001; 2001US-0277338P.  
PR 21-MAR-2001; 2001US-0277791P.  
PR 22-MAR-2001; 2001US-0277833P.  
PR 23-MAR-2001; 2001US-0278152P.  
PR 26-MAR-2001; 2001US-0278894P.  
PR 27-MAR-2001; 2001US-0278999P.  
PR 27-MAR-2001; 2001US-0279036P.  
PR 28-MAR-2001; 2001US-0279344P.  
PR 30-MAR-2001; 2001US-0279995P.  
PR 30-MAR-2001; 2001US-0280233P.  
PR 02-APR-2001; 2001US-0280802P.  
PR 02-APR-2001; 2001US-0280823P.  
PR 02-APR-2001; 2001US-0280900P.  
PR 04-APR-2001; 2001US-0281444P.  
PR 13-APR-2001; 2001US-0283675P.  
PR 30-APR-2001; 2001US-0287424P.  
PR 02-MAY-2001; 2001US-0288066P.  
PR 03-MAY-2001; 2001US-0288342P.  
PR 03-MAY-2001; 2001US-0288538P.  
PR 15-MAY-2001; 2001US-0291190P.  
PR 16-MAY-2001; 2001US-0291099P.  
PR 16-MAY-2001; 2001US-0291240P.  
PR 30-MAY-2001; 2001US-0294485P.  
PR 31-MAY-2001; 2001US-0294889P.  
PR 31-MAY-2001; 2001US-0294899P.  
PR 18-JUN-2001; 2001US-0299027P.  
PR 19-JUN-2001; 2001US-0299303P.  
PR 19-JUN-2001; 2001US-0299310P.  
PR 10-JUL-2001; 2001US-0304354P.  
PR 31-JUL-2001; 2001US-0309198P.  
PR 16-AUG-2001; 2001US-0312903P.  
PR 10-SEP-2001; 2001US-0318462P.  
PR 12-SEP-2001; 2001US-0318770P.  
PR 27-SEP-2001; 2001US-0325430P.  
PR 27-SEP-2001; 2001US-0325681P.  
PR 18-OCT-2001; 2001US-0330380P.

PR 31-OCT-2001; 2001US-0335301P.  
PR 14-NOV-2001; 2001US-0332172P.  
PR 14-NOV-2001; 2001US-0332271P.  
PR 14-NOV-2001; 2001US-0332272P.  
PR 14-NOV-2001; 2001US-0333184P.  
PR 14-NOV-2001; 2001US-0333272P.  
PR 21-NOV-2001; 2001US-0332094P.  
PR 03-DEC-2001; 2001US-0337426P.  
PR 03-DEC-2001; 2001US-0338092P.  
PR 04-DEC-2001; 2001US-0337185P.  
PR 03-JAN-2002; 2002US-0345705P.  
XX  
XX (PADI/) PADIGARU M.  
PA (SPYT/) SPYTEK K A.  
PA (SHEN/) SHENOY S G.  
PA (TAUP/) TAUPIER R J.  
PA (PENA/) PENA C E A.  
PA (LILL/) LI L.  
PA (ZERH/) ZERHUSEN B D.  
PA (GUSE/) GUSEV V Y.  
PA (JIWV/) JI W.  
PA (GORM/) GORMAN L.  
PA (MILL/) MILLER C E.  
PA (KEKU/) KEKUDA R.  
PA (PATT/) PATTURAJAN M.  
PA (GANG/) GANGOLLI E A.  
PA (VERN/) VERNET C A M.  
PA (GUOX/) GUO X S.  
PA (TCHE/) TCHERNEV V T.  
PA (FERN/) FERNANDES E R.  
PA (CASM/) CASMAN S J.  
PA (MALX/) MALYANKAR U M.  
PA (GERL/) GERLACH V.  
PA (LIUY/) LIU Y.  
PA (ANDE/) ANDERSON D W.  
PA (SPAD/) SPADERNA S K.  
PA (CATT/) CATTERTON E.  
PA (LEIT/) LEITE M W.  
PA (ZHON/) ZHONG H.  
PA (ALSO/) ALSOBROOK J P.  
PA (LEPL/) LEPLEY D M.  
PA (RIEG/) RIEGER D K.  
PA (BURG/) BURGESS C E.  
XX  
XX Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;  
PI Zehrusen BD, Gusev VY, Ji W, Gorman L, Miller CE, Kekuda R;  
PI Patturajan M, Gangolli EA, Vernet CAM, Guo XS, Tchernev VT;  
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;  
PI Anderson DW, Spaderna SK, Catterton E, Leite MW, Zhong H;  
PI Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;  
XX  
XX WPI; 2004-225693/21.  
DR N-PSDB; ADN61776.  
XX  
XX New NOVX polypeptides and nucleic acid molecules useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. cancer, diabetes,  
PT infection or obesity, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
XX Claim 1; SEQ ID NO 46; 786pp; English.  
XX  
XX The invention relates to an isolated polypeptide (designated NOVX, or  
CC NOV1-NOV127) comprising a sequence selected from 178 fully defined amino  
CC acid sequences (and their mature forms, variants and fragments). Also  
CC included are an isolated nucleic acid molecule encoding NOVX, a vector  
CC comprising the nucleic acid, a cell comprising the vector, methods for  
CC determining the presence or amount of the polypeptide or the nucleic acid  
CC molecule in a sample, methods for determining the presence of or  
CC predisposition to a disease associated with altered levels of expression  
CC of the above polypeptide or nucleic acid molecule in a first mammalian  
CC subject, a method for identifying an agent that binds to the above  
CC polypeptide, a method for identifying a potential therapeutic agent for  
CC use in the treatment of a pathology that is related to aberrant





XX OS Homo sapiens.  
 XX WO2003025168-A1.  
 XX PN 27-MAR-2003.  
 XX PD  
 XX PF 13-SEP-2002; 2002WO-JP009444.  
 XX PR 17-SEP-2001; 2001JP-00281992.  
 XX PR 02-OCT-2001; 2001JP-00306873.  
 XX PR 16-APR-2002; 2002JP-00113279.  
 XX PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX PI Nakanishi A, Uno Y, Sagiya Y;  
 XX WPI: 2003-313352/30.  
 XX DR N-PSDB; ACC79538, ACC79558.  
 XX PT Protein TCH169 with dicarboxylate transport activity for treatment and  
 XX diagnosis of diseases including liver disease, cancer and circulatory  
 XX disorders.  
 XX PS Claim 1; Fig 1-2; 132pp; Japanese.  
 XX CC The present invention describes protein TCH169 and its salts having  
 CC dicarboxylate transport activity. TCH169 has hepatotropic, cytostatic,  
 CC nephrotoxic, vasotropic and antidiabetic activities. The TCH169 protein  
 CC and polynucleotide can be used in the treatment, prevention and diagnosis  
 CC of liver disease (such as hepatitis, hepatic sclerosis and alcohol-  
 CC related liver disease); prostate disease (such as prostatitis and  
 CC prostatic hypertrophy); spleen disease (such as spleen hyperactivity);  
 CC kidney disease (such as nephritis, kidney failure, nephritis, dropsy and  
 CC diabetes-associated renal disease); metabolic disease (such as diabetes);  
 CC circulatory disease (such as hyperlipaemia and arteriosclerosis); and  
 CC cancer (such as non-small cell lung cancer, liver cancer, renal cancer,  
 CC ovarian cancer, prostate cancer, stomach cancer, pancreatic cancer,  
 CC breast cancer, colon cancer, bladder cancer and womb cancer). The present  
 CC sequence represents human TCH169, from the present invention  
 XX SQ Sequence 568 AA;

Query Match 45.6%; Score 1411.5; DB 6; Length 568;  
 Best Local Similarity 47.1%; Pred. No. 8.3e-139;  
 Matches 286; Conservative 100; Mismatches 150; Indels 71; Gaps 13;

Qy 4 LAALAKKWSARRLLVLLVPLALLPILFALPKPKGRCGLYVILLMAVTVCTEALPLSVTA 63  
 Db 1 MASALSYYVSKFVFILFVTLPLLLPLVILMPAKFVRCAYVILLMAIYVCTEVIPLAVTS 60  
 Qy 64 LLPILPFPFMGILPSSKVCPOVFLDTNFLSGLIMASAIERNLHRRIALKVLMLGVQ 123  
 Db 61 LMPVLLPFLQLLQSRQVCVQMKDNLFLGGLIVAVAVERNLHRRIALKVLMLGVQAK 120  
 Qy 124 PARLILGMVTVTSFLSMWLSNTASTAMPLIASAILKSLFGQDTRKDLPREGSDSTA 183  
 Db 121 PARLMLGFMVGTALLSMWISNTATTAMVPVIVEALIQM-----EATSAA 165  
 Qy 184 RGNGLRTVPTMQFLASSEGGAEDVPEAPLELPDD-----SKEEHRNINWKGFL 233  
 Db 166 TEAGLELV-----DKGKAK-----ELPGSQVIFEGPTLQQQDQBRKRLCKAMT 209  
 Qy 234 ISIPYSASIGGTATLTGAPNLILGLQSKSPQC-DVNFSGSWIFAPFLMLLFLVGM 292  
 Db 210 LCICVAASIGGTATLTGPNVLLGQNEUFLPDSKDLNVNFASFAPFAPNNVLLFAW 269  
 Qy 293 LWISPLYGMSW-RGW-----RKNSKLQDVAEDKAKAVIQEENFONLGPDKFAEQVFL 346  
 Db 270 LWLQFVYRNFNPKSWGCGLESKNEK-----ALKVLQEEYRKLGLPSFAINVLIC 322  
 Qy 347 FCLFAILLFSRDPKFIQWASL-FAPG---FVSDAVTGVAVITLFFPPSQKPSLKMWD 402  
 Db 347 FCLFAILLFSRDPKFIQWASL-FAPG---FVSDAVTGVAVITLFFPPSQKPSLKMWD 402

Db 323 FFLVILWFSRDPGMPGWLTVANVEGETKVKVSDATVAIFVATLLFIYPSQPK-----FN 378  
 Qy 403 FKAPNSETE-----PLLSWKKAQETVPMNIIILLGGGFAMAKGCEGSLGQQLH 455  
 Db 379 FRSQTEBERKTPFYPPPLLDKMKVQKVPWGLVLLGGGFALAKGSEASGLSVWNGQME 438  
 Qy 456 PLEHVPPLLAVLLITVVIATPTEFASNTATIIIFLPVLAELAIRLHVHPLYIMIGTVSC 515  
 Db 439 PLHAVPPAAITLILSLVAVFECTSNVATTTFLFPIFASMSRSIGLNPPLYIMLPTLSA 498  
 Qy 516 SYAFMLPVSTPNSTAFSTGHLLVKDMVTRTGLLNMGMVLLLSLAMNTWAOAIFOLGTFP 575  
 Db 499 SFAPMLPVATPNAIVFTYGHVKVADMVKTGVIMNIIGVFCVFLAVNTWGRAIFOLDHPP 558  
 Qy 576 DWAN-TH 581  
 Db 559 DWANVTH 565

RESULT 13  
 AAE38764  
 ID AAE38764 standard; protein; 568 AA.  
 XX  
 AC AAE38764;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Human 69624 protein.  
 XX  
 KW Human; 69624; transporter protein; neurological disorder; therapy;  
 KW atherosclerosis; cardiac hypertrophy; ischaemia reperfusion injury;  
 KW metastatic disorder; haematopoietic neoplastic disorder; leukaemia;  
 KW arthritis; multiple sclerosis; encephalomyelitis; myasthenia gravis;  
 KW carcinoma; cell proliferation; autoimmune disorder; diabetes mellitus;  
 KW renal disorder; colon; hepatic disorder; hypocitraturia; calcium stone;  
 KW mental retardation; Canavan disease; differentiative disorder; sarcoma;  
 KW systemic lupus erythematosus; cardiovascular disorder; arteriosclerosis;  
 KW atrial fibrillation; forensic identification; pain.  
 XX  
 OS Homo sapiens.  
 XX FH  
 Key Location/Qualifiers  
 Domain 6..554  
 FT /note= "Sodium sulphate symporter domain"  
 Domain 14..31  
 FT /note= "Transmembrane domain"  
 Domain 41..167  
 FT /note= "Cotransporter sodium/dicarboxylate domain"  
 Domain 53..73  
 FT /note= "Transmembrane domain"  
 Domain 87..103  
 FT /note= "Transmembrane domain"  
 Modified-site 118..123  
 FT /note= "N-myristoylation site"  
 Domain 124..140  
 FT /note= "Transmembrane domain"  
 Domain 202..375  
 FT /note= "Cotransporter sodium/dicarboxylate domain"  
 Domain 210..234  
 FT /note= "Transmembrane domain"  
 Modified-site 226..231  
 FT /note= "N-myristoylation site"  
 Domain 253..272  
 FT /note= "Transmembrane domain"  
 Modified-site 291..293  
 FT /note= "Protein kinase C phosphorylation site"  
 Domain 309..331  
 FT /note= "Transmembrane domain"  
 Domain 353..370  
 FT /note= "Transmembrane domain"  
 Modified-site 372..374  
 FT /note= "Protein kinase C phosphorylation site"  
 Modified-site 381..384







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:36:57 ; Search time 23.069 Seconds  
(without alignments)  
1941.542 Million cell updates/sec

Title: US-10-017-479A-5  
Perfect score: 3094  
Sequence: 1 MAALAAKAKWSARRLLVL.....HAANVTALPPALTNNVTQTL 600

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1333.5	43.1	599	4	US-09-949-016-9866
2	1208.5	39.1	601	4	US-09-949-016-9977
3	1188	38.4	627	4	US-09-556-916-2
4	1186	38.3	627	4	US-09-949-016-6840
5	1183.5	38.3	626	4	US-09-556-916-14
6	1182	38.2	627	4	US-09-556-916-4
7	1180	38.1	627	4	US-09-556-916-8
8	1177.5	38.1	626	4	US-09-556-916-16
9	1175.5	38.0	626	4	US-09-556-916-20
10	1174	37.9	627	4	US-09-556-916-10
11	1169.5	37.8	626	4	US-09-556-916-22
12	1019	32.9	581	4	US-09-556-916-6
13	1014.5	32.8	580	4	US-09-556-916-18
14	1011	32.7	581	4	US-09-556-916-12
15	1006.5	32.5	580	4	US-09-556-916-24
16	950	30.7	561	4	US-09-949-016-8161
17	813	26.3	233	4	US-09-489-847-176
18	740	23.9	527	4	US-09-602-787A-516
19	517	16.7	368	4	US-09-270-767-42029
20	501	16.2	470	4	US-09-543-681A-5952
21	461.5	14.9	335	4	US-09-602-787A-518
22	392.5	12.7	524	3	US-09-134-001C-5457
23	308.5	10.0	180	4	US-09-270-767-42669
24	307.5	9.9	421	4	US-09-248-796A-20749
25	265.5	8.6	230	4	US-09-270-767-43713
26	255.5	8.3	194	4	US-09-270-767-59102
27	251	8.1	132	4	US-09-270-767-57987

28	234	7.6	547	4	US-09-489-039A-13843	Sequence 13843, A
29	207.5	6.7	624	4	US-09-543-681A-4343	Sequence 4343, Ap
30	206	6.7	169	4	US-09-270-767-57286	Sequence 57286, A
31	204	6.6	494	4	US-09-543-681A-7033	Sequence 7033, Ap
32	203	6.6	302	4	US-09-902-540-10445	Sequence 10445, A
33	189.5	6.1	614	4	US-09-489-039A-12605	Sequence 12605, A
34	189	6.1	596	4	US-09-902-540-13547	Sequence 13547, A
35	185	6.0	548	4	US-09-902-540-11870	Sequence 11870, A
36	167	5.4	694	4	US-09-252-991A-22637	Sequence 22637, A
37	164.5	5.3	470	4	US-09-438-185A-209	Sequence 209, App
38	161.5	5.2	493	4	US-09-540-236-2120	Sequence 2120, Ap
39	157	5.1	443	4	US-09-602-787A-532	Sequence 532, App
40	142.5	4.6	430	3	US-09-134-001C-2981	Sequence 2981, Ap
41	141	4.6	478	3	US-09-134-001C-4637	Sequence 4637, Ap
42	137	4.4	441	4	US-09-489-039A-10612	Sequence 10612, A
43	130.5	4.2	597	4	US-09-489-039A-14112	Sequence 14112, A
44	122	3.9	167	4	US-09-248-796A-20740	Sequence 20740, A
45	120.5	3.9	363	4	US-09-270-767-42613	Sequence 42613, A
46	115.5	3.7	319	4	US-09-270-767-57927	Sequence 57927, A
47	114	3.7	429	4	US-08-311-731A-287	Sequence 287, App
48	113	3.7	437	4	US-09-543-681A-6984	Sequence 6984, Ap
49	113	3.7	1042	4	US-09-252-991A-30444	Sequence 30444, A
50	112.5	3.6	1077	4	US-09-412-210-1	Sequence 1, Appli
51	112.5	3.6	1077	4	US-10-121-911A-1	Sequence 1, Appli
52	111	3.6	211	4	US-09-198-452A-1166	Sequence 1166, Ap
53	111	3.6	381	4	US-09-710-279-3284	Sequence 3284, Ap
54	110.5	3.6	3421	4	US-09-452-638-53	Sequence 53, Appli
55	110	3.6	494	4	US-09-603-208A-232	Sequence 232, App
56	108.5	3.5	449	4	US-08-328-352-7512	Sequence 7512, Ap
57	108.5	3.5	554	4	US-09-543-681A-5774	Sequence 5774, Ap
58	108	3.5	509	4	US-09-134-000C-5949	Sequence 5949, Ap
59	108	3.5	611	2	US-08-677-049-2	Sequence 2, Appli
60	106.5	3.4	580	4	US-09-657-252-2	Sequence 2, Appli
61	106	3.4	540	3	US-09-433-994-2	Sequence 2, Appli
62	105	3.4	574	4	US-08-902-540-10370	Sequence 10370, A
63	104.5	3.4	456	3	US-09-058-389A-2	Sequence 2, Appli
64	104.5	3.4	456	3	US-09-611-781-2	Sequence 2, Appli
65	104.5	3.4	457	4	US-09-949-016-7211	Sequence 7211, Ap
66	104.5	3.4	482	4	US-09-328-352-7784	Sequence 7784, Ap
67	104.5	3.4	1912	4	US-09-495-714C-2	Sequence 2, Appli
68	104.5	3.4	1977	4	US-08-495-714C-4	Sequence 4, Appli
69	104	3.4	502	4	US-09-252-991A-23327	Sequence 23327, A
70	104	3.4	535	4	US-09-107-433-3342	Sequence 3342, Ap
71	104	3.4	540	4	US-09-583-110-3359	Sequence 3359, Ap
72	104	3.4	549	4	US-09-489-039A-13779	Sequence 13779, A
73	103	3.3	115	4	US-09-602-787A-520	Sequence 520, App
74	103	3.3	345	4	US-09-902-540-9836	Sequence 9836, Ap
75	103	3.3	514	4	US-09-543-681A-5352	Sequence 5352, Ap
76	103	3.3	594	4	US-09-252-991A-27335	Sequence 27335, A
77	102.5	3.3	598	4	US-09-252-991A-28599	Sequence 28599, A
78	102	3.3	505	4	US-09-328-352-7470	Sequence 7470, Ap
79	102	3.3	1014	4	US-09-252-991A-29868	Sequence 29868, A
80	101	3.3	395	4	US-09-543-681A-6203	Sequence 6203, Ap
81	100.5	3.2	562	4	US-09-489-039A-8574	Sequence 8574, Ap
82	100	3.2	344	4	US-08-311-731A-172	Sequence 172, App
83	100	3.2	518	4	US-09-252-991A-30478	Sequence 30478, A
84	100	3.2	532	4	US-09-252-991A-22421	Sequence 22421, A
85	100	3.2	693	4	US-09-949-016-9666	Sequence 9666, Ap
86	99.5	3.2	279	4	US-09-602-787A-658	Sequence 658, App
87	99.5	3.2	326	3	US-09-058-389A-3	Sequence 3, Appli
88	99.5	3.2	326	3	US-09-611-781-3	Sequence 3, Appli
89	99.5	3.2	408	3	US-09-134-001C-3410	Sequence 3410, Ap
90	99.5	3.2	463	4	US-09-583-110-5223	Sequence 5223, Ap
91	99.5	3.2	878	4	US-09-438-185A-729	Sequence 729, App
92	99	3.2	534	4	US-09-252-991A-30725	Sequence 30725, A
93	98.5	3.2	300	4	US-09-107-532A-5624	Sequence 5624, Ap
94	98.5	3.2	313	4	US-09-902-540-15660	Sequence 15660, A
95	98.5	3.2	378	4	US-09-055-097-1	Sequence 1, Appli
96	98.5	3.2	378	4	US-09-373-902-1	Sequence 1, Appli
97	98.5	3.2	393	4	US-09-949-016-11567	Sequence 11567, A
98	98	3.2	525	4	US-09-902-540-15103	Sequence 15103, A
99	98	3.2	767	4	US-09-540-236-2346	Sequence 2346, Ap
100	97.5	3.2	378	4	US-09-831-630-13	Sequence 13, Appli

## ALIGNMENTS

## RESULT 1

```
US-09-949-016-9866
; Sequence 9866, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9866
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9866

Query Match      43.1%; Score 1333.5; DB 4; Length 599;
Best Local Similarity 45.4%; Pred. No. 9.1e-126;
Matches 280; Conservative 109; Mismatches 175; Indels 53; Gaps 13;

Qy      4 LAALAKKWSARRLLVLLPLALLPLFALPKPEGKCLYVILLMAYVWCTEALPLSVTA 63
Db      8 MATCQALWAVRSYLVFFVPIILLPLILVPSKEAYCAVAIILMAFWCTEALPLAVTA 67

Qy      64 LLPILFPFGIILPSSKVCQYFDLTNFLSLGLINASAEERNLHRRIRIALKVLMLVG 123
Db      68 LFPILFPFMGIVDASEVAYEYKDSNLLFPFGLLVAVIAVEHNLHRRIRIALRVLLV 127

Qy      124 PARILGMVTTSLMNLNTASTAMMLPIASAILKSLFGQDRDKLDPREGEDSTA 183
Db      128 PAPILGFMVLTAFSLWISNTATSNMVPPIAHVLDQL-----HSSQASSNVE 176

Qy      184 RGNGLRTV-----PTFMQFLASSEGHAEDVEAPLELPDDSKEEHRRNIWKGF 233
Db      177 EGSNNPTFELQEPSPOKEVTKLDNGQA-----LPVTSASSEGRAHLSQKHLHLT 228

Qy      234 -ISIPYSASIGGTATLTGTAPNLILGLQKSPFPQCDVNVFGSWFIAPFLMLLLV 291
Db      229 SLVCYSASIGGIATLTGTAPNLVLQGINSLFPQNGNVNVFASWFSFAPPTWVILL 288

Qy      292 WLWISFLYGGMSWRGWRKKN----SKLDVAEDKAKAVIQEENLGPRIKFAEQAVFI 347
Db      289 WLWQILFLGNFR----KNFGICEKWE-QQQAYCVIQTEHLLGPMTFEAKAISILF 343

Qy      348 CLFAILLFSRDPKFIQWASLFAF-----GPVSDAVTGVAIVTILFPFPKPSLKW 402
Db      344 VILVLLWFTREPGFVLGNLAFNAKGSMSVSGTVAIFGIIMFIIPSKFGLTQ--D 401

Qy      403 FKPNSETP-----LLSWKKAQETVPMNIIILGGGFAMAKGCEESGLSAGIQLH 460
Db      402 PENFGKLPAGLGLDWTQNMKNPWNIVLLGGGYALAKGERSGLSEWLNKUTPLQSV 461

Qy      461 PLLAVLLITVVIATFFEFASNTATIIIFLPVLAELAIRLHVHPLYLMIPTGVSCSY 520
Db      462 PAPAIAILSLVATFECTSNVATITIFLPILASMAQAICLHPLYVYMLPCTLATSLA 521

Qy      521 LPVSTPPNSIAFSGHLLVKDMVTRTGLLMNLMGLVLLSLAMNTWAQAI FOLGTFPD 580
Db      522 LPVATPPNAIVFSFGDLKVLDMARAGFLNIIIGVLIILALAINSGIPLFSLHSF 580
```

## RESULT 2

```
US-09-949-016-9977
; Sequence 9977, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9977
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9977

Query Match      39.1%; Score 1208.5; DB 4; Length 601;
Best Local Similarity 40.8%; Pred. No. 4e-113;
Matches 250; Conservative 119; Mismatches 186; Indels 57; Gaps 12;

Qy      15 RRLVLLVPLALLPLFALPKPEGKCLYVILLMAYVWCTEALPLSVTALLPLPFP 74
Db      17 RRFLFVFTVLLPLPLIVLTKAECAVTLFVATFWLTALPLSVTALLPSLMLPFG 76

Qy      75 ILPSSKVCQYFDLTNFLSLGLINASAEERNLHRRIRIALKVLMLVGQPARLILGM 134
Db      77 IMPSKVASAYPKDFHLLIIGVICLATSIEKNLHRRIRIALKVLMLVGVNPAWLT 136

Qy      135 TFLSMLNTASTAMMLPIASAILKSLFG-----QRDTRKDLPRGEDESTA 186
Db      137 TAFSLMNLNTASTAMVMPDIAEAVVQIINAEAEVATQMTYFNGSTNHGLEIDES 196

Qy      187 GL-----RTVPTMQFLASSEGHAED---VEAPLELPDDSKEEHRRNIWKGF 233
Db      197 EINERKEKTPV-----GYNNDTGKISKVELEKNSGMRTKYRT-KKHVTRKLT 246

Qy      234 -ISIPYSASIGGTATLTGTAPNLILGLQKSPFPQCDVNVFGSWFIAPFLMLLLV 292
Db      247 CLCIAVSSITGGITTTIGTSTNLIIPAEYFNTRYPCDCLNFGSWFTSFPAALI 306

Qy      293 LMTISFLYGGMSWRGWRKKNSKLDVAEDKAKAVIQEENLGPRIKFAEQAVFI 352
Db      307 IWLQWFLFNGFR--MFKCGKTKVQKCAEVIKQEQYKLGPIRYQEIIVTLVLF 365

Qy      353 LLFSRDPKFIQWASLFAF-----GPVSDAVTGVAIVTILFPFPKPSLKW 410
Db      366 LWFSDRPGFVPGNSALFSEYPGFATDS--VALLIGLLFLIIPAKTLTK----- 419

Qy      411 -----BPLLSWKKQAQETVPMNIIILGGGFAMAKGCEESGLSAGIQLH 465
Db      420 VADYSEPLITWKEFQSFMPMDIALLVGGGFALADGCEESGLSKWIGKSLPGL 479

Qy      466 VLIITVVIATFFEFASNTATIIIFLPVLAELAIRLHVHPLYLMIPTGVSCSY 525
Db      480 ILISLWVTSLSLTVASNPNATITILFPLTSLPABAIHVNPPLYIILIPSTLC 539

Qy      526 PPNISAFSGHLLVKDMVTRTGLLMNLMGLVLLSLAMNTWAQAI FOLGTFPD 585
Db      527 LPVATPPNAIVFSFGDLKVLDMARAGFLNIIIGVLIILALAINSGIPLFSLHS 585
```

Db 540 PPNALVPSYGLKVIDMVKAGLVNIVGVAVMLGICTWIVPMFDLYTPSWA----- 592  
 QY 586 TALPPALTNNTV 597  
 Db 593 ----PAMSNETM 600

RESULT 3

US-09-556-916-2

; Sequence 2, Application US/09556916

; Patent No. 6548271

; GENERAL INFORMATION:

; APPLICANT: Turner, Alex

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Nehls, Michael

; APPLICANT: Friedrich, Glenn

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins

; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916

; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FaastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-556-916-2

Query Match 38.4%; Score 1188; DB 4; Length 627;

Best Local Similarity 38.6%; Pred. No. 5e-111;

Matches 241; Conservative 127; Mismatches 194; Indels 62; Gaps 12;

QY 15 RRLVLLVPLALLPLFALPPKGRCLYVILLMAVYCTEALPLSVTALLPIILPFPNG 74  
 Db 11 RKLWVVCVPLLLPLPVLPHPSEASCAYLVITAVYVSEAVPLGAAALVPAFLYPPFG 70

QY 75 ILPSSKVCQYFLDTNPLFSLGLMASAIEERNLHRRALKVLMVGVQPARLILGMVT 134  
 Db 71 VLRSNEVAEYFKNTLLVGVICVAAVEKNLHRRALKVLMVGVQPARLILGMVT 130

QY 135 TSFLSNLSTASTAMMLPIASAILKSLF-----GQDTRKDLF----- 173  
 Db 131 TLLSNLSTNTTAWMPIVEAVLQELVSAEDQLVAGNSNTEAEPISLDVKNQSPSL 190

QY 174 -----REGDSTAARGNGLRTPV--TEMQFLASSEGG--HAEDVEAPLELPDDSK 220  
 Db 191 ELIFVNEEDRSNADLTLLMHNENLVGPSITNPIKTANQHQKKQHPSEKQPVLTPSPRK 250

QY 221 EE-----EHRNINWKGFLISIPYSASIGGTATLTGTAPNLLLGQLKSPFPQCDVNF 273  
 Db 251 QKLNRYRSHHDOMICKCLSLISYSATIGLTTIIGTSTLSIFLEHFNNOYPAEAVNF 310

QY 274 GSWFIFAPPLMLLLVGLWLTISFLYGMNSWRGW----RKKNKSLQDVAEDKAKAVIOEE 329  
 Db 311 GTWFLFSFPLISLMLVSNFWMHLFLGCFNFKETCSLSKKKTKREQLSEKR----IQEE 366

QY 330 FQNLGPIKFAEQAVFLFCLFALLFSRDPKIPGWSLIF-APGFVSDAVTGVAVITILF 388  
 Db 367 YEKLGDISYPMVMTGFFFLMTVLMFTREPGVPGWDSFFFEKKGYRTDATVSVFLGFLF 426

QY 389 FPPSQKSLKMWDFDK--APNSE-----TEPLLSWKKAEQETVPWNIIILLGGGFAMAKGE 442  
 Db 427 LIPAKKPC-----FGKNDGENQEHSLGTEPIITWKDFQKTMPEWIVLVGGYALASGK 482

QY 443 ESGLSAWIGQHLPLEHPVPLAVLITVIAFFTEFASNTATIIIFLPVLAELAIRLHV 502  
 Db 483 SSGLSWIGNQMLSSLSLPPWAVTLLACILVSVTEFVSNPATITIFILPILCSLSETHI 542

QY 503 HPLYLMIPGTVSCSYAFMLPVSTPPNSIAFSTGHLLVKDMVTRTGLLMLMGVLLSLAMN 562  
 Db 543 NPLYTLIPVTMCISFAVNLPGVNPNAIVFSYGHQCIKDMVKAGLVNIVGLIVMVAIN 602

QY 563 TWAQAIQFOLGTTPDPWANTHAANVT 586  
 Db 603 TWGVSFLHLDTPAWA--RVSNIT 624

RESULT 4

US-09-949-016-6840

; Sequence 6840, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FaastSeq for Windows Version 4.0

; SEQ ID NO 6840

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Human

US-09-949-016-6840

Query Match 38.3%; Score 1186; DB 4; Length 627;

Best Local Similarity 38.5%; Pred. No. 8e-111;

Matches 240; Conservative 128; Mismatches 194; Indels 62; Gaps 12;

QY 15 RRLVLLVPLALLPLFALPPKGRCLYVILLMAVYCTEALPLSVTALLPIILPFPNG 74  
 Db 11 RKLWVVCVPLLLPLPVLPHPSEASCAYLVITAVYVSEAVPLGAAALVPAFLYPPFG 70

QY 75 ILPSSKVCQYFLDTNPLFSLGLMASAIEERNLHRRALKVLMVGVQPARLILGMVT 134  
 Db 71 VLRSNEVAEYFKNTLLVGVICVAAVEKNLHRRALKVLMVGVQPARLILGMVT 130

QY 135 TSFLSNLSTASTAMMLPIASAILKSLF-----GQDTRKDLF----- 173  
 Db 131 TLLSNLSTNTTAWMPIVEAVLQELVSAEDQLVAGNSNTEAEPISLDVKNQSPSL 190

QY 174 -----REGDSTAARGNGLRTPV--TEMQFLASSEGG--HAEDVEAPLELPDDSK 220  
 Db 191 ELIFVNEEDRSNADLTLLMHNENLVGPSITNPIKTANQHQKKQHPSEKQPVLTPSPRK 250

QY 221 EE-----EHRNINWKGFLISIPYSASIGGTATLTGTAPNLLLGQLKSPFPQCDVNF 273  
 Db 251 QKLNRYRSHHDOMICKCLSLISYSATIGLTTIIGTSTLSIFLEHFNNOYPAEAVNF 310

QY 274 GSWFIFAPPLMLLLVGLWLTISFLYGMNSWRGW----RKKNKSLQDVAEDKAKAVIOEE 329  
 Db 311 GTWFLFSFPLISLMLVSNFWMHLFLGCFNFKETCSLSKKKTKREQLSEKR----IQEE 366

QY 330 FQNLGPIKFAEQAVFLFCLFALLFSRDPKIPGWSLIF-APGFVSDAVTGVAVITILF 388  
 Db 367 YEKLGDISYPMVMTGFFFLMTVLMFTREPGVPGWDSFFFEKKGYRTDATVSVFLGFLF 426

QY 389 FPPSQKSLKMWDFDK--APNSE-----TEPLLSWKKAEQETVPWNIIILLGGGFAMAKGE 442  
 Db 427 LIPAKKPC-----FGKNDGENQEHSLGTEPIITWKDFQKTMPEWIVLVGGYALASGK 482

QY 443 ESGLSAWIGQHLPLEHPVPLAVLITVIAFFTEFASNTATIIIFLPVLAELAIRLHV 502  
 Db 483 SSGLSWIGNQMLSSLSLPPWAVTLLACILVSVTEFVSNPATITIFILPILCSLSETHI 542

QY 503 HPLYLMIPGTVSCSYAFMLPVSTPPNSIAFSTGHLLVKDMVTRTGLLMLMGVLLSLAMN 562  
 Db 543 NPLYTLIPVTMCISFAVNLPGVNPNAIVFSYGHQCIKDMVKAGLVNIVGLIVMVAIN 602

```
QY 563 TWAQAFOLGTFPDWANTHAANT 586
Db 603 TWGVSFLHLDTPAWA--RVSNT 624

RESULT 5
US-09-556-916-14
; Sequence 14, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-14

Query Match 38.3%; Score 1183.5; DB 4; Length 626;
Best Local Similarity 38.5%; Pred. No. 1.4e-110;
Matches 240; Conservative 127; Mismatches 195; Indels 61; Gaps 12;

QY 15 RRLVLLVPLALLPILFALPPKRGCLYVILLMAYVWCTEALPLSVTALLPILPPFMG 74
Db 11 RKLVLVVCVPLLLPLPLVLPSSSEASCAYVLIVTAVYVWSEAVPLGAAALVPAFLYPPFG 70
QY 75 ILPSKVCPOYFLDTNFIPLSLGIMASAEIRNLHRIALKVLMLVGVOPARILGMVT 134
Db 71 VLRSNEVAEYFQNTLLVGVICVAARVKNLHKRIALRMVLMAKAGPMGLLCPMCC 130
QY 135 TSFLSMLSNTASTAMMLPIASAILKSLF-----GORDTRKDL----- 173
Db 131 TTLSSMLSNTSTTAWMPIVEAVLQELVSAEDQVAGNSNTEEAEPISLDVKNSPSL 190
QY 174 -----REGEDSTAARVGNGLRTP--TEMQFLASSEGG---HAEDVEAPLELPDDSK 221
Db 191 ELIFVNEESNADLTLMHNNENLNGVPSITNPITKANQHQKKQHPSEKQPVLTPSPRKQ 250
QY 222 E-----EHRNNTWKGLISIPYSASIGGTATLTGTAPNLILGLKSPFPQCDVVNFG 274
Db 251 KLNKRYRSHDDQMLCKLSLSISYSATIGGLTTIIGTSTLSIFLEHNNQYPAAEVNFG 310
QY 275 SWFIPAPPLMLLFLLVGMLNITFLYGGMSRWG----RKNSKLQDVADKAKAVIQEEF 330
Db 311 TWFLSPFISLIMLVSWFWHMLFLGCKFKETCSLSKKKTKREQLSEK-----IQEEY 366
QY 331 QNLGPIKFAEQAVFILCLFAILLFSDPKFIPGWASLF-APGFVSDAVTGVAIVTILFF 389
Db 367 EKLGDISYPEMVTGFFFTILMTVLTREPGVPGWDSFFEKKGYRTDATVSVFLGFLFL 426
QY 390 FPSQKPSLKNWFDFK--APNSE-----TEPLLSSWKAQETVPWNIIILGGGFAMAKGCEE 443
Db 427 IPAKKPC-----FGKNDGENQEHSLGTEPIITWKDFQKTPWEIVLVGGYALASGKS 482
QY 444 SGLSAWIGQQLHPLHVPPLLAVALITVIAFFTEPASNTATITIFLPVLAELAIRLHV 503
Db 483 SGLSTWIGNQMLSLSSLPWAVTLLACILVSVTEFVSNPATITIFLPILCSLSETHIN 542
QY 504 PLYLMIPTGVSCSYAFMLPVSTPPNSAFSTGHLLVKDMVKTGLLNMGLVLLSLAMNT 563
Db 543 PLYTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVGNVIGLIVMVAINT 602
QY 564 WAQAFOLGTFPDWANTHAANT 586
```

```
Db 603 WGVSLFHLDTYPAWA--RVSNT 623

RESULT 6
US-09-556-916-4
; Sequence 4, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-4

Query Match 38.2%; Score 1182; DB 4; Length 627;
Best Local Similarity 38.5%; Pred. No. 2e-110;
Matches 240; Conservative 127; Mismatches 195; Indels 62; Gaps 12;

QY 15 RRLVLLVPLALLPILFALPPKRGCLYVILLMAYVWCTEALPLSVTALLPILPPFMG 74
Db 11 RKLVLVVCVPLLLPLPLVLPSSSEASCAYVLIVTAVYVWSEAVPLGAAALVPAFLYPPFG 70
QY 75 ILPSKVCPOYFLDTNFIPLSLGIMASAEIRNLHRIALKVLMLVGVOPARILGMVT 134
Db 71 VLRSNEVAEYFQNTLLVGVICVAARVKNLHKRIALRMVLMAKAGPMGLLCPMCC 130
QY 135 TSFLSMLSNTASTAMMLPIASAILKSLF-----GORDTRKDL----- 173
Db 131 TTLSSMLSNTSTTAWMPIVEAVLQELVSAEDQVAGNSNTEEAEPISLDVKNSPSL 190
QY 174 -----REGEDSTAARVGNGLRTP--TEMQFLASSEGG---HAEDVEAPLELPDDSK 220
Db 191 ELIFVNEESNADLTLMHNNENLNGVPSITNPITKANQHQKKQHPSEKQPVLTPSPRK 250
QY 221 E-----EHRNNTWKGLISIPYSASIGGTATLTGTAPNLILGLKSPFPQCDVVNFG 273
Db 251 OKLNKRYRSHDDQMLCKLSLSISYSATIGGLTTIIGTSTLSIFLEHNNQYPAAEVNFG 310
QY 274 GSWFIPAPPLMLLFLLVGMLNITFLYGGMSRWG----RKNSKLQDVADKAKAVIQEE 329
Db 311 GTWFLSPFISLIMLVSWFWHMLFLGCKFKETCSLSKKKTKREQLSEK-----IQEE 366
QY 330 QNLGPIKFAEQAVFILCLFAILLFSDPKFIPGWASLF-APGFVSDAVTGVAIVTILF 388
Db 367 YEKLGDISYPEMVTGFFFTILMTVLTREPGVPGWDSFFEKKGYRTDATVSVFLGFLFL 426
QY 389 FPSQKPSLKNWFDFK--APNSE-----TEPLLSSWKAQETVPWNIIILGGGFAMAKGCEE 442
Db 427 LIPAKKPC-----FGKNDGENQEHSLGTEPIITWKDFQKTPWEIVLVGGYALASGSK 482
QY 443 ESSLASWIGQQLHPLHVPPLLAVALITVIAFFTEPASNTATITIFLPVLAELAIRLHV 502
Db 483 SGLSTWIGNQMLSLSSLPWAVTLLACILVSVTEFVSNPATITIFLPILCSLSETHI 542
QY 503 HPLYLMIPTGVSCSYAFMLPVSTPPNSAFSTGHLLVKDMVKTGLLNMGLVLLSLAMN 562
Db 543 NPLYTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVGNVIGLIVMVAIN 602
QY 563 TWAQAFOLGTFPDWANTHAANT 586
Db 603 TWGVSFLHLDTPAWA--RVSNT 624
```

## RESULT 8

US-09-556-916-16  
; Sequence 16, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 626  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-556-916-16

Query Match 38.1%; Score 1177.5; DB 4; Length 626;  
Best Local Similarity 38.4%; Pred. No. 5.8e-110;  
Matches 239; Conservative 127; Mismatches 196; Indels 61; Gaps 12;

Qy	15	RLLLVLLVPLALLPILFALPKGRCGLVILLMAVYVCTEALPLSVTALLPILPPFPG 74
Db	11	RKLLVVCVPLLLPLVLPVHPSEASCAVLIVAVYVSEAVPLGAALVPAFLYPPFG 70
Qy	75	ILPSSKVCQYFLDTNFLFSLGLMASAIEERNLHRRIALKVLMLVGVQPARLILGMVT 134
Db	71	VLSRNEVAEYFKNTLLVGVICVAAAEKWNLHRIALRMVLMAGAKPGMLLLCFMCC 130
Qy	135	TSFSLMWLNTASTAMMLPIASAILKSLF-----GQDTRKDL- 173
Db	131	TLLSMLWLSNTSTTAMVPIVEAVLQELVSADEQLVAGNSNTEAEPLSLDVKNQSPSL 190
Qy	174	-----REGDSTAAVRNGRLTVP--TEMQFLASSEGG--HAEDVEAPLELPDDSK 221
Db	191	ELIFVNEEDSNADLTLLMNEENLNGVPSITNPKTANQHGKQHPQSKPQVLTPSPRK 250
Qy	222	E-----EHRNIIWKGLFISIPYSASIGGTATLTGTAPNLIILGQLKSPFPQCDVNF 274
Db	251	KLNRKYRSHDQMICKCLSLSLISYSATIGLTTIGTSTLSIFLEHFNQYPAAEVNF 310
Qy	275	GSWFIPAPLMLLLVGLWLTISFLYGGMSWRGW-----RKNSKLQDVAEDKAKAVIOEE 330
Db	311	GTWFLFSPISILMLVSVFWMHFLGNCFKETCSLSKKKTKREQLSEKR-----IQEE 366
Qy	331	QNLGPIKFAEQAVFILFCLPAILLFSRDPKFI PGWASLF-APGFVSDAVTGVAVITILF 389
Db	367	EKLGDISYPMVMTGFFFLMTVLMFTREPGVPGWDSFEKKGYRTDATSVFGLFLF 426
Qy	390	FPSQKPSLKNWFDK--APNSE-----TEPLLSWKAQSTVPWNIIILGGGFAMAKGCE 443
Db	427	IPAKKPC-----FGKNNGENQHSLSGTESIITWKDFQKTMPEWIVILVGGYALASGSK 482
Qy	444	ESGLSAWIGQQLHLEHPVPLLAVALITVVIAPFTFASNTATITIFLPVLAELAIRLHV 503
Db	483	SSGLSTWIGNQMLSSLPWAVTLLACILVSIVTEFVSNPATITIFLPILCSLSETLHI 542
Qy	504	HPYLMIPQTVSCSAFAMLPVSTPPNSIAFSGHLLVKDMVTRTGLLMLMGVLLSLANNT 563
Db	543	NPLYTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNVI GLVIMVAIN 602
Qy	564	TWQAIFOLGTFPPDWANTHAANVT 586
Db	603	TWVSLFHLDTYPAWA--RVSNIT 624

## RESULT 9

US-09-556-916-20

## RESULT 7

US-09-556-916-8  
; Sequence 8, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 627  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-556-916-8

Query Match 38.1%; Score 1180; DB 4; Length 627;  
Best Local Similarity 38.5%; Pred. No. 3.2e-110;  
Matches 240; Conservative 127; Mismatches 195; Indels 62; Gaps 12;

Qy	15	RLLLVLLVPLALLPILFALPKGRCGLVILLMAVYVCTEALPLSVTALLPILPPFPG 74
Db	11	RKLLVVCVPLLLPLVLPVHPSEASCAVLIVAVYVSEAVPLGAALVPAFLYPPFG 70
Qy	75	ILPSSKVCQYFLDTNFLFSLGLMASAIEERNLHRRIALKVLMLVGVQPARLILGMVT 134
Db	71	VLSRNEVAEYFKNTLLVGVICVAAAEKWNLHRIALRMVLMAGAKPGMLLLCFMCC 130
Qy	135	TSFSLMWLNTASTAMMLPIASAILKSLF-----GQDTRKDL- 173
Db	131	TLLSMLWLSNTSTTAMVPIVEAVLQELVSADEQLVAGNSNTEAEPLSLDVKNQSPSL 190
Qy	174	-----REGDSTAAVRNGRLTVP--TEMQFLASSEGG--HAEDVEAPLELPDDSK 220
Db	191	ELIFVNEEDSNADLTLLMNEENLNGVPSITNPKTANQHGKQHPQSKPQVLTPSPRK 250
Qy	221	EE-----EHRNIIWKGLFISIPYSASIGGTATLTGTAPNLIILGQLKSPFPQCDVNF 273
Db	251	QXLRKYRSHDQMICKCLSLSLISYSATIGLTTIGTSTLSIFLEHFNQYPAAEVNF 310
Qy	274	GSWFIPAPLMLLLVGLWLTISFLYGGMSWRGW-----RKNSKLQDVAEDKAKAVIOEE 329
Db	311	GTWFLFSPISILMLVSVFWMHFLGNCFKETCSLSKKKTKREQLSEKR-----IQEE 366
Qy	330	QNLGPIKFAEQAVFILFCLPAILLFSRDPKFI PGWASLF-APGFVSDAVTGVAVITILF 388
Db	367	EKLGDISYPMVMTGFFFLMTVLMFTREPGVPGWDSFEKKGYRTDATSVFGLFLF 426
Qy	389	FPSQKPSLKNWFDK--APNSE-----TEPLLSWKAQSTVPWNIIILGGGFAMAKGCE 442
Db	427	LIPAKKPC-----FGKNNGENQHSLSGTESIITWKDFQKTMPEWIVILVGGYALASGSK 482
Qy	443	ESGLSAWIGQQLHLEHPVPLLAVALITVVIAPFTFASNTATITIFLPVLAELAIRLHV 502
Db	483	SSGLSTWIGNQMLSSLPWAVTLLACILVSIVTEFVSNPATITIFLPILCSLSETLHI 542
Qy	503	HPYLMIPQTVSCSAFAMLPVSTPPNSIAFSGHLLVKDMVTRTGLLMLMGVLLSLANNT 562
Db	543	NPLYTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNVI GLVIMVAIN 602
Qy	563	TWQAIFOLGTFPPDWANTHAANVT 586
Db	603	TWVSLFHLDTYPAWA--RVSNIT 624

```
; Sequence 20, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-20

Query Match      38.0%; Score 1175.5; DB 4; Length 626;
Best Local Similarity 38.4%; Pred. No. 9.2e-110;
Matches 239; Conservative 127; Mismatches 196; Indels 61; Gaps 12;

QY 15 RRLVLLVPLALLPLFALPKPEGRCGLYVILLMAVYWCTEALPLSVTALLPIILPFFMG 74
Db 11 RKLLVVCVPLLLPLPVLHPSSSEACAYVLIIVTAVYVSEAVPLGAAALVPALPFFPG 70
QY 75 ILPSSKVCPOYFLDTNLFSLGLIMASAEERNLHRRIALKVLMLVGQVOPARILGMVMT 134
Db 71 VLRSNEVAAYFKNTLLLVGVICVAAAVEKNLHKRIALRMVLMAKAGPKMLLCFMCC 130
QY 135 TSFLSMWLSNTASTAMMLPIASAILKSLF-----GQDRTRKDLPL----- 173
Db 131 TTLLSMWLSNTSTTAMVMPIVEAVLQELVSAEDQLVAGNSNTEEAEPISLDVKNQPSL 190
QY 174 -----REGEDSTAAVRGNGLRTP--TEMQFLASSEG--HAEDVEAPLELPDDSK 221
Db 191 ELIFVNESNADLTTLMHENENLNGVPSITNPITANOHQKQHPSEKQVLTSPRKQ 250
QY 222 E-----EHRNRNWKFLSIPYSASIGGTATLTGTAPNLLILGOLKSPFQCDVNVF 274
Db 251 KLNKRYRSHDQMICKLSLSISYSATIGLTIIGTSTSLIFLEHFNNOYPAAEVNFG 310
QY 275 SWTFAPALMLLLVLGVMLWISFLYGMWSRWG-----RKNSKLQDVADKAKAVIOEE 330
Db 311 TWFLFSPFISLIMLVSWFMWHLFLGCFNKETCSLSKKKTKRQLSEKR-----IQEE 366
QY 331 QNLGPIKFAEQVFLFCFLFALLFSRDPKPIPGWASLF-APGFVSDAVTGVAVITLFF 389
Db 367 EKLGDISYPEWVTGFFILMTVLFTRFPGVPGWDSFFEKKGYRTDATVSVFLGFLFL 426
QY 390 FPSQKPSLKWDFDK--APNSE-----TEPLLSSWKAQETVPMNILLGGGFAMAKGCE 443
Db 427 IPAKKPC-----FGKNKGNGENQHSLSGTESIITWKDFQKTMPEWIVILVGGGYALASGKS 482
QY 444 SGLSANTGGQLHPLHVPPLAVLLITVIAFFTEPASNTATIIIFLPVLAELAIRLHVH 503
Db 483 SGLSTWIGNQMLSSSLPFWAVTLLACILSVITFEVSNPATITIFLPILCSLSETHIN 542
QY 504 PLYLMIPTVSCSYAFMLPVSTPNSTAFSTGHLVVKDMVRTGILLMNLGMVLLSLAMNT 563
Db 543 PLYTLIPVTCISFAVMLPVGNPNNAIVFSYGHQIKDMVKAGLVGNVIGLVIMVAINT 602
QY 564 WQAQIFOLGTFPPDWTANHAANVT 586
Db 603 WGVSLFHLDTYPAWA--RVSNTIT 623

RESULT 10
US-09-556-916-10
; Sequence 10, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-10

Query Match      37.9%; Score 1174; DB 4; Length 627;
Best Local Similarity 38.3%; Pred. No. 1.3e-109;
Matches 239; Conservative 127; Mismatches 196; Indels 62; Gaps 12;

QY 15 RRLVLLVPLALLPLFALPKPEGRCGLYVILLMAVYWCTEALPLSVTALLPIILPFFMG 74
Db 11 RKLLVVCVPLLLPLPVLHPSSSEACAYVLIIVTAVYVSEAVPLGAAALVPALPFFPG 70
QY 75 ILPSSKVCPOYFLDTNLFSLGLIMASAEERNLHRRIALKVLMLVGQVOPARILGMVMT 134
Db 71 VLRSNEVAAYFKNTLLLVGVICVAAAVEKNLHKRIALRMVLMAKAGPKMLLCFMCC 130
QY 135 TSFLSMWLSNTASTAMMLPIASAILKSLF-----GQDRTRKDLPL----- 173
Db 131 TTLLSMWLSNTSTTAMVMPIVEAVLQELVSAEDQLVAGNSNTEEAEPISLDVKNQPSL 190
QY 174 -----REGEDSTAAVRGNGLRTP--TEMQFLASSEG--HAEDVEAPLELPDDSK 220
Db 191 ELIFVNESNADLTTLMHENENLNGVPSITNPITANOHQKQHPSEKQVLTSPRK 250
QY 221 EE-----EHRNRNWKFLSIPYSASIGGTATLTGTAPNLLILGOLKSPFQCDVNVF 273
Db 251 QLNKRYRSHDQMICKLSLSISYSATIGLTIIGTSTSLIFLEHFNNOYPAAEVNFG 310
QY 274 GWFIFAPFAPMLLLVLGVMLWISFLYGMWSRWG-----RKNSKLQDVADKAKAVIOEE 329
Db 311 GTWFLFSPFISLIMLVSWFMWHLFLGCFNKETCSLSKKKTKRQLSEKR-----IQEE 366
QY 330 FQNLGPIKFAEQVFLFCFLFALLFSRDPKPIPGWASLF-APGFVSDAVTGVAVITLFF 388
Db 367 YEKLGDISYPEWVTGFFILMTVLFTRFPGVPGWDSFFEKKGYRTDATVSVFLGFLFL 426
QY 389 FPSQKPSLKWDFDK--APNSE-----TEPLLSSWKAQETVPMNILLGGGFAMAKGCE 442
Db 427 LIPAKKPC-----FGKNKGNGENQHSLSGTESIITWKDFQKTMPEWIVILVGGGYALASGKS 482
QY 443 EGSLSANTGGQLHPLHVPPLAVLLITVIAFFTEPASNTATIIIFLPVLAELAIRLHVH 502
Db 483 SSGLSWTWIGNQMLSSSLPFWAVTLLACILSVITFEVSNPATITIFLPILCSLSETHI 542
QY 503 HPLYLMIPTVSCSYAFMLPVSTPNSTAFSTGHLVVKDMVRTGILLMNLGMVLLSLAMN 562
Db 543 NPLYTLIPVTCISFAVMLPVGNPNNAIVFSYGHQIKDMVKAGLVGNVIGLVIMVAIN 602
QY 563 TWAQAI FOLGTFPPDWTANHAANVT 586
Db 603 TWGVSFLHLDTPAWA--RVSNTIT 624

RESULT 11
US-09-556-916-22
; Sequence 22, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
```

; APPLICANT: Zambrowicz, Brian		; APPLICANT: Friedrich, Glenn	
; APPLICANT: Nehls, Michael		; APPLICANT: Sands, Arthur T.	
; FILE OF INVENTION: No. 6548271el Human Transporter Proteins		; FILE OF INVENTION: No. 6548271el Human Transporter Proteins	
; CURRENT APPLICATION NUMBER: US/09/556,916		; CURRENT APPLICATION NUMBER: US/09/556,916	
; CURRENT FILING DATE: 2000-04-21		; CURRENT FILING DATE: 2000-04-21	
; NUMBER OF SEQ ID NOS: 32		; NUMBER OF SEQ ID NOS: 32	
; SOFTWARE: FastSeq for Windows Version 3.0		; SOFTWARE: FastSeq for Windows Version 3.0	
; SEQ ID NO 22		; SEQ ID NO 6	
; LENGTH: 626		; LENGTH: 581	
; TYPE: PRT		; TYPE: PRT	
; ORGANISM: Homo sapiens		; ORGANISM: Homo sapiens	
US-09-556-916-22		US-09-556-916-6	
Query Match		Query Match	
Best Local Similarity		Best Local Similarity	
Matches 238; Conservative 127; Mismatches 197; Indels 61; Gaps 12;		Matches 217; Conservative 120; Mismatches 179; Indels 108; Gaps 14;	
QY	15 RRLVLLVLPALLPILFALPKGRCLYVILLMAVYCTEALPLSVTALLPIILPFPFG 74	QY	15 RRLVLLVLPALLPILFALPKGRCLYVILLMAVYCTEALPLSVTALLPIILPFPFG 74
Db	11 RKLLVVCVPLLLPLVLPHPSEASCAVYLVITAVYVYSEAVPLGAAALVPAFLYPPFG 70	Db	11 RKLLVVCVPLLLPLVLPHPSEASCAVYLVITAVYVYSEAVPLGAAALVPAFLYPPFG 70
QY	75 ILPSKVCQYFDTNPLFLSGLIMASAEERNLHRRIALKVLMLVGVPARLILGMVT 134	QY	75 ILPSKVCQYFDTNPLFLSGLIMASAEERNLHRRIALKVLMLVGVPARLILGMVT 134
Db	71 VLRSNEVAEYFKNTLLLVGVICVAAAEKWNLHKRIALRMVLMAGAKPGMLLLCFMCC 130	Db	71 VLRSNEVAEYFKNTLLLVGVICVAAAEKWNLHKRIALRMVLMAGAKPGMLLLCFMCC 130
QY	135 TSFLSMLSNTASTAMLPISAILKSLF-----GQDTRKDLF----- 173	QY	135 TSFLSMLSNTASTAMLPISAILKSLF-----GQDTRKDLF----- 173
Db	131 TTLSSMLSNTSTTAMVPIEAVLQELVSAEDQLVAGNSNTEAEPISLDVKNQSPSL 190	Db	131 TTLSSMLSNTSTTAMVPIEAVLQELVSAEDQLVAGNSNTEAEPISLDVKNQSPSL 190
QY	174 -----REGEDSTAARVNGRLTVP--TEMQFLASSEG--HAEDVEAPLELPDDSK 221	QY	174 -----REGEDSTAARVNGRLTVP--TEMQFLASSEG--HAEDVEAPLELPDDSK 220
Db	191 ELIFVNEESNADLTTLMHENLNGVPSITNPIKTANQHQKQHPQSEKPVLTSPRKQ 250	Db	191 ELIFVNEEDRSNADLTTLMHENLNGVPSITNPIKTANQHQKQHPQSEKPVLTSPRK 250
QY	222 E-----EHRNINWGFILISIPYASIGGTATLTGTAPNLILLGOLKSPFPQCDVNF 274	QY	221 EE-----EHRNINWGFILISIPYASIGGTATLTGTAPNLILLGOLKSPFPQCDVNF 273
Db	251 KLNRYRSHHDMICKLSLSISYSATIGLTTIIGTSTSLIFLHFNNQYPAABVNVFG 310	Db	251 QKLNRYRSHHDMICKLSLSISYSATIGLTTIIGTSTSLIFLHFNNQYPAABVNVNF 310
QY	275 SWFIFAPFLMLFLVGLWISFLYGGMSWRGW-----RKNSKLQDVAEDKAKAVIOEF 330	QY	274 GSWFIFAPFLMLFLVGLWISFLYGGMSWRGW-----RKNSKLQDVAEDKAKAVIOBE 329
Db	311 TWFLSFPSILIMLVSWFMHWFPGCNFKETCSLKKKTKRQLSEKR-----IQEY 366	Db	311 GTWFLSFPSILIMLVSWFMHWFPGCNFKETCSLKKKTKRQLSEKR-----IQEE 366
QY	331 QNLGPIKAEQAVFIFCLFALLFRDPKPIFGWASLF--APGFVSDAVTGVAIVTILFF 389	QY	330 FQNLGPIKAEQAVFIFCLFALLFRDPKPIFGWASLF--APGFVSDAVTGVAIVTILF 388
Db	367 EKLGDISYPMVGTGFFILMTVLTREPGFVPGWDSFEKKGYRTDATSVFLGFLFL 426	Db	367 YELGDISYPMVGTGFFILMTVLTREPGFVPGWDSFEKKGYRTDATSVFLGFLFL 426
QY	390 FPSQKSLKMWPDFK--APNSE-----TEPLLSWKKKAQETVPWNIILLGGGFAMAKCEE 443	QY	389 FPSQKSLKMWPDFK--APNSE-----TEPLLSWKKKAQETVPWNIILLGGGFAMAKCEE 442
Db	427 IPAKKPC-----FGKNDGENQHSLSGTESIITWKDFQKTPMPEIVILVGGGVALASGKS 482	Db	427 LIPAKKPC-----FGKNDGENQHSLSGTESIITWKDFQKTPMPEIVILVGGGVALASGK 482
QY	444 SGLSAMIWGQLHPLHVPPLAVLLITVVIAPFTEFASNTATIIIFLPVLAELAIRLHV 503	QY	443 BSGLSAMIWGQLHPLHVPPLAVLLITVVIAPFTEFASNTATIIIFLPVLAELAIRLHV 502
Db	483 SGLSTWIGNQMLSSLPWPWATLLACILSVIVTEFVSNPATITIFLPILCSLSETOHIN 542	Db	483 SSGLSTWIGNQMLSSLPWPWATLLACILSVIVTEFVSNPATITIFLPIIL----- 533
QY	504 PLYLMPGTVCSSYAFMLPVSTPPNSIAFSTGHLLVKDMVYRTGLLMLNMGVLLSLAMNT 563	QY	503 HPLYLMPGTVCSSYAFMLPVSTPPNSIAFSTGHLLVKDMVYRTGLLMLNMGVLLSLAMN 562
Db	543 PLYTLIPVTCISFAPVLPVGNPPNVAIVFSYGHCHQIKDMVKAGLVNVLGLVIVNVAINT 602	Db	534 -----CS-----LVKAGLVNVLGLVIVNVAIN 556
QY	564 WQAQIFQLGTFFPDWANTHAANVT 586	QY	563 TWAQAIQLGTFFPDWANTHAANVT 586
Db	603 WGVSLFHLDTYPAWA--RVSNT 623	Db	557 TWGVSLFHLDTYPAWA--RVSNT 578
RESULT 12		RESULT 13	
US-09-556-916-6		US-09-556-916-18	
; Sequence 6, Application US/09556916		; Sequence 18, Application US/09556916	
; Patent No. 6548271		; Patent No. 6548271	
; GENERAL INFORMATION:		; GENERAL INFORMATION:	
; APPLICANT: Turner, Alex		; APPLICANT: Turner, Alex	
; APPLICANT: Zambrowicz, Brian		; APPLICANT: Zambrowicz, Brian	
; APPLICANT: Nehls, Michael		; APPLICANT: Nehls, Michael	
; APPLICANT: Friedrich, Glenn		; APPLICANT: Friedrich, Glenn	
; APPLICANT: Sands, Arthur T.		; APPLICANT: Sands, Arthur T.	

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 580  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-556-916-18

Query Match 32.8%; Score 1014.5; DB 4; Length 580;  
Best Local Similarity 34.7%; Pred. No. 1.6e-93;  
Matches 216; Conservative 120; Mismatches 180; Indels 107; Gaps 14;

QY 15 RLLVLLVPLALLPILFALPKPEGRCLVYVILLMAVYVCTEALPLSVLTALLPILPFFPMG 74  
DB 11 RLLVLLVCPVLLLLPLVLPHPSEASCAYLVITAVYVWSEAVPLGAAALVPAFLYPPFG 70  
QY 75 ILPSKVCPOYFLDTNLFSLGLIMASAEERNLHRRIALKVLMLVGVQPARILGMVMT 134  
DB 71 VLRSNEVAAYEFKNTLLVGVICVAAAVEKNLHRRIALRMVLMAGAKPGMLLLCFMCC 130  
QY 135 TSFLSMWLSNTASTAMMLPIASAILKSLF-----GQDTRKDL----- 173  
DB 131 TTLSMWLSNTSTTAMVPIVEAVLQELVSAEDQVLVAGNSNTEEAEPISLVKNSQPSL 190  
QY 174 -----REGEDSTAARVGNGLRTP--TEMQFLASSEG--HAEDVEAPLELPDDSK 220  
DB 191 ELIFVNEEDSNADLTLMENENLNGVPSITNPKTANQHQKQHPQSEKQVLTSPRK 250  
QY 222 E-----EHRNRTWKGLFISIPYSASIGTATLTGTAPNLLILGQLKSFPPQCDVNF 274  
DB 251 KLNRYRSHHDQMKCLSLISYSATIGLTTIGTSTLSLIFLEHFNNOYPAAEVNVF 310  
QY 275 SWETPAPLMFLVGLWMLTSLYLGMSWRGW----RKNSKLDQVADKAKAVIOEEF 330  
DB 311 TWLFSFPISLIMLVSWFWMHLLFCNFKETCSLKKKTKRQJSEKR-----IQEEY 366  
QY 331 QNLGPIKFAEQAVFLFCLFAILLFSDPKPIPGWASLF-APGFVSDAVTGVAVITLFF 399  
DB 367 EKLGDISPEVMTGFFFLMTVLTREPGVPGWDSFFEKKGYRTDATVSVFLGFLFL 426  
QY 390 FPSQKPSLKWDFK--APNSE-----TEPLLSWKAQETVPWNILLLGGFPAKACGE 443  
DB 427 IPAKKPC----FGKNDGENQEHSLGTEPIITWKDFQKTPMPEIIVILVGGGYALASGKS 482  
QY 444 SGLSAGWIGQLHLEHPVPLLAIVLITVIAFFTEFASNTATIIIFLPVLAELAIRLHV 503  
DB 483 SGLSTWIGNQMLSSLPWAVTLACILVSVTEFVSNPATITIFLPIL----- 532  
QY 504 PLYLMIPTGVSCSYAFMLPVSTPPNSIAFSTGHLLVVKDMVRTGLLMLNMGVLLSLAMNT 563  
DB 533 -----CS-----LVKAGLVNVVIGLVVMAINT 556  
QY 564 WQAQIFQLGTFPDWANTHAANVT 586  
DB 557 WGVSLFHLDTYPAWA--RVSNIT 577

RESULT 14  
US-09-556-916-12  
; Sequence 12, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 12  
; LENGTH: 581  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-556-916-12

Query Match 32.7%; Score 1011; DB 4; Length 581;  
Best Local Similarity 34.6%; Pred. No. 3.5e-93;  
Matches 216; Conservative 120; Mismatches 180; Indels 108; Gaps 14;

QY 15 RLLVLLVPLALLPILFALPKPEGRCLVYVILLMAVYVCTEALPLSVLTALLPILPFFPMG 74  
DB 11 RLLVLLVCPVLLLLPLVLPHPSEASCAYLVITAVYVWSEAVPLGAAALVPAFLYPPFG 70  
QY 75 ILPSKVCPOYFLDTNLFSLGLIMASAEERNLHRRIALKVLMLVGVQPARILGMVMT 134  
DB 71 VLRSNEVAAYEFKNTLLVGVICVAAAVEKNLHRRIALRMVLMAGAKPGMLLLCFMCC 130  
QY 135 TSFLSMWLSNTASTAMMLPIASAILKSLF-----GQDTRKDL----- 173  
DB 131 TTLSMWLSNTSTTAMVPIVEAVLQELVSAEDQVLVAGNSNTEEAEPISLVKNSQPSL 190  
QY 174 -----REGEDSTAARVGNGLRTP--TEMQFLASSEG--HAEDVEAPLELPDDSK 220  
DB 191 ELIFVNEEDSNADLTLMENENLNGVPSITNPKTANQHQKQHPQSEKQVLTSPRK 250  
QY 221 E-----EHRNRTWKGLFISIPYSASIGTATLTGTAPNLLILGQLKSFPPQCDVNF 273  
DB 251 QLNRYRSHHDQMKCLSLISYSATIGLTTIGTSTLSLIFLEHFNNOYPAAEVNVF 310  
QY 274 GSWETPAPLMFLVGLWMLTSLYLGMSWRGW----RKNSKLDQVADKAKAVIOEE 329  
DB 311 GTWFLFSFPISLIMLVSWFWMHLLFCNFKETCSLKKKTKRQJSEKR-----IQEE 366  
QY 330 FQNLGPIKFAEQAVFLFCLFAILLFSDPKPIPGWASLF-APGFVSDAVTGVAVITLFF 388  
DB 367 YEKLGDISPEVMTGFFFLMTVLTREPGVPGWDSFFEKKGYRTDATVSVFLGFLFL 426  
QY 389 FPSQKPSLKWDFK--APNSE-----TEPLLSWKAQETVPWNILLLGGFPAKACGE 442  
DB 427 LIPAKKPC----FGKNDGENQEHSLGTEPIITWKDFQKTPMPEIIVILVGGGYALASGSK 482  
QY 443 EGSLSAGWIGQLHLEHPVPLLAIVLITVIAFFTEFASNTATIIIFLPVLAELAIRLHV 502  
DB 483 SGLSTWIGNQMLSSLPWAVTLACILVSVTEFVSNPATITIFLPIL----- 533  
QY 503 HPLYLMIPTGVSCSYAFMLPVSTPPNSIAFSTGHLLVVKDMVRTGLLMLNMGVLLSLAMN 562  
DB 534 -----CS-----LVKAGLVNVVIGLVVMAIN 556  
QY 563 WQAQIFQLGTFPDWANTHAANVT 586  
DB 557 WGVSLFHLDTYPAWA--RVSNIT 578

RESULT 15  
US-09-556-916-24  
; Sequence 24, Application US/09556916  
; Patent No. 6548271  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Alex  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins  
; FILE REFERENCE: 8535-041-999  
; CURRENT APPLICATION NUMBER: US/09/556,916  
; CURRENT FILING DATE: 2000-04-21

```
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-556-916-24

Query Match      32.5%; Score 1006.5; DB 4; Length 580;
Best Local Similarity 34.5%; Pred. No. 1e-92;
Matches 215; Conservative 120; Mismatches 181; Indels 107; Gaps 14;

QY 15 RLLVLLVPLALLPLFALPKEGECYVILLMAYVWCTEALPLSVTALLPIILPPFMG 74
Db 11 RKLLVVCVPLLLPLVLPVHPSEASCAYLVITVAVYVSEAVPLGAAALVPAFLYPPFG 70

QY 75 ILPSSKVCQYFLDTNFLFLGLIMASAJEERNLHRRIALKVLMLVGVOVARLIILGMVT 134
Db 71 VLRSNEVAEYFNKNTLLLVGVICVAAVEKNLHRIALRMVLMAGAKPGMLLCFMCC 130

QY 135 TSFLSNMLSNASTAMMLPIASAILKSLF-----GQDTRKDLF----- 173
Db 131 TTLLSNMLSNSTTAMVMPIVEAVLQELVSAEDEQLVAGNSNTEEAEPISLDVKNQSPSL 190

QY 174 -----REGEDSTAARNGRLRTP--TEMQFLASSEGG---HADDVEAPLELPDDSKB 221
Db 191 ELIFVNEESNADLTILMHENENLNGVPSITNPITANQHOKQKHPSEQEPQVLTSPRRQ 250

QY 222 E-----EHRNINWKGFLISIPYSASIGGTATLTCTAPNLILLGOLKSPFPQCDVNVFG 274
Db 251 KLNRYRSHHDQMICKUSLSISYSATIGLTIIGTSTSLFLEHFNNOYPAAEVNVFG 310

QY 275 SWFIAPFLMLFLVGLWGLWISFLYGGMSWRGW----RKNSKLQDVADKAKAVIQBEF 330
Db 311 TWFLSFPSILIMLVSVNFWHMLFLGCFKETSLSKKKTKRQLSEK-----IQEY 366

QY 331 QNLGPIKFAEQAVFILFCLFAILLFSRDPKPIPGWASLP-APGFVSDAVTGVAIVTILFF 389
Db 367 EKLGDISYPEMVTGFFFIILMTVLFETREPFGVPGWDSFPEKGYRTDATSVFLGLLFL 426

QY 390 FPSQPSLKWWFDFK--APNSE---TEPLLSKKAOETVPWNIIILLGGGFAMAKGCEE 443
Db 427 IPAKKPC----FGKNDGENQEHSLGTESIITWKDFQKTPMWEIVILVGGGYALASGSKS 482

QY 444 SGLSAWIGGQLHPLEHVPLLAFLITVVIAPFTEFASNTATIIIFLPVLAELAIRLHVH 503
Db 483 SGLSTWIGNQMLSSLSLPWAVTILLACILVSVITFEVSNPATITIFLPIL----- 532

QY 504 PLYLMIPGTVSCSYAFMLPVSTPPNSIAFSTGHLLVKDMVRTGLLNMNLGMVLLSLAMNT 563
Db 533 -----CS-----LVKAGLGWVIGLVIVNVAINT 556

QY 564 WQAIFQLGTFPDWANTHAANVT 586
Db 557 WGVSLFLDLYPAWA--RVSNIT 577
```

Search completed: June 30, 2005, 09:02:19  
Job time : 26.069 secs

**THIS PAGE IS BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:59:01 ; Search time 77.8579 Seconds  
(without alignments)  
2973.073 Million cell updates/sec

Title: US-10-017-479a-5

Perfect score: 3094

Sequence: 1 MAALAAAKVNSARRLLVL.....HAANYTALPPALTNNVTQL 600

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1722976 seqs, 385795295 residues

Total number of hits satisfying chosen parameters: 1722976

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3094	100.0	600	17	US-10-718-359-14
2	3092	99.9	600	14	US-10-167-994-13
3	2891	87.0	602	14	US-10-167-994-4
4	2326.5	75.2	520	15	US-10-108-260A-3862
5	1412.5	45.7	616	15	US-10-092-900A-46
6	1412.5	45.7	616	15	US-10-092-900A-48
7	1412.5	45.7	616	15	US-10-403-161-68
8	1411.5	45.6	568	13	US-10-173-519-2
9	1411.5	45.6	568	13	US-10-403-161-72
10	1411.5	45.6	568	16	US-10-490-080-1
11	1411.5	45.6	568	17	US-10-718-359-6
					Sequence 14, Appl
					Sequence 13, Appl
					Sequence 4, Appl
					Sequence 3862, Ap
					Sequence 46, Appl
					Sequence 48, Appl
					Sequence 68, Appl
					Sequence 2, Appl
					Sequence 7, Appl
					Sequence 1, Appl
					Sequence 6, Appl

Sequence 149, App  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 44, Appl  
Sequence 70, Appl  
Sequence 12, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 5, Appl  
Sequence 12, Appl  
Sequence 13, Appl  
Sequence 6, Appl  
Sequence 4, Appl  
Sequence 23, Appl  
Sequence 3, Appl  
Sequence 11, Appl  
Sequence 50, Appl  
Sequence 74, Appl  
Sequence 118, App  
Sequence 5, Appl  
Sequence 9, Appl  
Sequence 2, Appl  
Sequence 7, Appl  
Sequence 14, Appl  
Sequence 8, Appl  
Sequence 16, Appl  
Sequence 20, Appl  
Sequence 10, Appl  
Sequence 22, Appl  
Sequence 52, Appl  
Sequence 76, Appl  
Sequence 6, Appl  
Sequence 18, Appl  
Sequence 12, Appl  
Sequence 24, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 8, Appl  
Sequence 1947, Ap  
Sequence 1949, Ap  
Sequence 346, App  
Sequence 10, Appl  
Sequence 176, App  
Sequence 436, App  
Sequence 127443,  
Sequence 160, App  
Sequence 127444,  
Sequence 516, App  
Sequence 3754, Ap  
Sequence 250656,  
Sequence 2022, Ap  
Sequence 254988,  
Sequence 4, Appl  
Sequence 354, App  
Sequence 354, App  
Sequence 354, App  
Sequence 2634, Ap  
Sequence 21334, A  
Sequence 518, App  
Sequence 8, Appl  
Sequence 7, Appl  
Sequence 1187, Ap  
Sequence 594, App  
Sequence 34252, A  
Sequence 28036, A  
Sequence 9458, Ap  
Sequence 980, App

85 369 11.9 234 15 US-10-335-977-5397 Sequence 5397, Ap  
 86 366 11.8 453 15 US-10-369-493-9936 Sequence 9936, Ap  
 87 360 11.6 457 15 US-10-369-493-4266 Sequence 4266, Ap  
 88 352 11.4 236 16 US-10-767-701-43427 Sequence 43427, A  
 89 346 11.2 1006 17 US-10-741-849-7011 Sequence 7011, Ap  
 90 345 11.2 163 15 US-10-424-599-184881 Sequence 184881,  
 91 320 10.3 174 16 US-10-425-115-335649 Sequence 335649,  
 92 318.5 10.3 499 14 US-10-238-075-500 Sequence 500, App  
 93 308.5 10.0 300 16 US-10-767-701-46317 Sequence 46317, A  
 94 296 9.6 162 11 US-09-833-245-1948 Sequence 1948, Ap  
 95 295.5 9.6 162 11 US-09-833-245-1948 Sequence 1948, Ap  
 96 283 9.1 487 15 US-10-369-493-23559 Sequence 23559, A  
 97 263.5 8.5 164 15 US-10-424-599-205152 Sequence 205152,  
 98 252 8.1 54 9 US-09-864-761-46701 Sequence 46701, A  
 99 207.5 6.7 487 15 US-10-369-493-749 Sequence 749, App  
 100 203 6.6 109 16 US-10-767-701-44228 Sequence 44228, A

# ALIGNMENTS

RESULT 1  
 US-10-718-359-14  
 ; Sequence 14, Application US/10718359  
 ; Publication No. US20050095240A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.  
 ; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION  
 ; FILE REFERENCE: 275.0008 0101  
 ; CURRENT APPLICATION NUMBER: US/10718,359  
 ; CURRENT FILING DATE: 2003-11-20  
 ; PRIOR APPLICATION NUMBER: 60/428,469  
 ; PRIOR FILING DATE: 2002-11-22  
 ; PRIOR APPLICATION NUMBER: 60/459,441  
 ; PRIOR FILING DATE: 2003-04-01  
 ; NUMBER OF SEQ ID NOS: 31  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 14  
 ; LENGTH: 600  
 ; TYPE: PRT  
 ; ORGANISM: rat Nadc3  
 US-10-718-359-14

Query Match 100.0%; Score 3094; DB 17; Length 600;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-273;  
 Matches 600; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAALAAKAKVWSARRLLVLLVPLALLPILFALPPKGRCLYVILLMAVYVWCTEALPLS 60  
 DB 1 MAALAAKAKVWSARRLLVLLVPLALLPILFALPPKGRCLYVILLMAVYVWCTEALPLS 60

QY 61 VTALLPILPPFMGILPSSKVCPOYFDTNPLFLSGLIMASAEERNLHRRALKVLMV 120  
 DB 61 VTALLPILPPFMGILPSSKVCPOYFDTNPLFLSGLIMASAEERNLHRRALKVLMV 120

QY 121 GVQPARLILGMVTTTSFLSMWLSNTASTAMMLPIASAILKSLFGQRTDKLPREGEDST 180  
 DB 121 GVQPARLILGMVTTTSFLSMWLSNTASTAMMLPIASAILKSLFGQRTDKLPREGEDST 180

QY 181 AAVRGNGLTRVPTQMFLASSEGGAEDVEAPLEPDDSKEEHRRNIWKGLISIPYSA 240  
 DB 181 AAVRGNGLTRVPTQMFLASSEGGAEDVEAPLEPDDSKEEHRRNIWKGLISIPYSA 240

QY 241 SIGGTATLTGTAPNLILGOLKSPFQCDVNVFGSWFIAPPLMLLLVGLWMLISFLYG 300  
 DB 241 SIGGTATLTGTAPNLILGOLKSPFQCDVNVFGSWFIAPPLMLLLVGLWMLISFLYG 300

QY 301 GMSWRGWRKKNKSLQDVAEDKAKAVIQEFGNLGPIKFAEQAVFILFCLFAILLFSRDPK 360  
 DB 301 GMSWRGWRKKNKSLQDVAEDKAKAVIQEFGNLGPIKFAEQAVFILFCLFAILLFSRDPK 360

QY 361 FIPGWASLFPAGFVSDAVTGVAVITLFFPFSQKPSLKWPFDFKAPNSETPEPLLSWKKAQ 420  
 DB 361 FIPGWASLFPAGFVSDAVTGVAVITLFFPFSQKPSLKWPFDFKAPNSETPEPLLSWKKAQ 420

Db 361 FIPGWASLFPAGFVSDAVTGVAVITLFFPFSQKPSLKWPFDFKAPNSETPEPLLSWKKAQ 420  
 QY 421 ETVPWNIIILLGGGFAMAKGCEESGLSAWIGQLHPLHVPPLLAVALITVVIAFTTEFA 480  
 Db 421 ETVPWNIIILLGGGFAMAKGCEESGLSAWIGQLHPLHVPPLLAVALITVVIAFTTEFA 480  
 QY 481 SNTATIIIFLPVLAELAIRLHVHPLYLMPGTVSCSYAFMLPVSTPPNSIAFSTGHLIVK 540  
 Db 481 SNTATIIIFLPVLAELAIRLHVHPLYLMPGTVSCSYAFMLPVSTPPNSIAFSTGHLIVK 540  
 QY 541 DMVRTGLLNMGMVLLLSLAMNTWAQAIQOLGTFPDWANTHAANTALPPALTNNVTQTL 600  
 Db 541 DMVRTGLLNMGMVLLLSLAMNTWAQAIQOLGTFPDWANTHAANTALPPALTNNVTQTL 600

RESULT 2  
 US-10-167-994-13  
 ; Sequence 13, Application US/10167994  
 ; Publication No. US20030082647A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reenan, Robert A.  
 ; APPLICANT: Rogina, Blanka  
 ; APPLICANT: Helfand, Stephen L.  
 ; TITLE OF INVENTION: TRANSPORTER PROTEIN  
 ; FILE REFERENCE: 13407-013001  
 ; CURRENT APPLICATION NUMBER: US/10/167,994  
 ; CURRENT FILING DATE: 2002-06-12  
 ; PRIOR APPLICATION NUMBER: US 60/255,013  
 ; PRIOR FILING DATE: 2000-12-12  
 ; PRIOR APPLICATION NUMBER: US 10/017,479  
 ; PRIOR FILING DATE: 2001-12-12  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13  
 ; LENGTH: 600  
 ; TYPE: PRT  
 ; ORGANISM: Rattus norvegicus  
 US-10-167-994-13

Query Match 99.9%; Score 3092; DB 14; Length 600;  
 Best Local Similarity 99.8%; Pred. No. 1.6e-273;  
 Matches 599; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAALAAKAKVWSARRLLVLLVPLALLPILFALPPKGRCLYVILLMAVYVWCTEALPLS 60  
 DB 1 MAALAAKAKVWSARRLLVLLVPLALLPILFALPPKGRCLYVILLMAVYVWCTEALPLS 60

QY 61 VTALLPILPPFMGILPSSKVCPOYFDTNPLFLSGLIMASAEERNLHRRALKVLMV 120  
 DB 61 VTALLPILPPFMGILPSSKVCPOYFDTNPLFLSGLIMASAEERNLHRRALKVLMV 120

QY 121 GVQPARLILGMVTTTSFLSMWLSNTASTAMMLPIASAILKSLFGQRTDKLPREGEDST 180  
 DB 121 GVQPARLILGMVTTTSFLSMWLSNTASTAMMLPIASAILKSLFGQRTDKLPREGEDST 180

QY 181 AAVRGNGLTRVPTQMFLASSEGGAEDVEAPLEPDDSKEEHRRNIWKGLISIPYSA 240  
 DB 181 AAVRGNGLTRVPTQMFLASSEGGAEDVEAPLEPDDSKEEHRRNIWKGLISIPYSA 240

QY 241 SIGGTATLTGTAPNLILGOLKSPFQCDVNVFGSWFIAPPLMLLLVGLWMLISFLYG 300  
 DB 241 SIGGTATLTGTAPNLILGOLKSPFQCDVNVFGSWFIAPPLMLLLVGLWMLISFLYG 300

QY 301 GMSWRGWRKKNKSLQDVAEDKAKAVIQEFGNLGPIKFAEQAVFILFCLFAILLFSRDPK 360  
 DB 301 GMSWRGWRKKNKSLQDVAEDKAKAVIQEFGNLGPIKFAEQAVFILFCLFAILLFSRDPK 360

QY 361 FIPGWASLFPAGFVSDAVTGVAVITLFFPFSQKPSLKWPFDFKAPNSETPEPLLSWKKAQ 420  
 DB 361 FIPGWASLFPAGFVSDAVTGVAVITLFFPFSQKPSLKWPFDFKAPNSETPEPLLSWKKAQ 420

QY 421 ETVPWNIIILLGGGFAMAKGCEESGLSAWIGQLHPLHVPPLLAVALITVVIAFTTEFA 480  
 DB 421 ETVPWNIIILLGGGFAMAKGCEESGLSAWIGQLHPLHVPPLLAVALITVVIAFTTEFA 480

Db 421 ETVPNWIIILLGGGFAMAKCESGLSAGWIGQLHPLHVPBLLAVLLITVVIAFTTEFA 480  
QY 481 SNTATIIIFLPVLAELAIRLHVHPLYLMIPTGVSCSYAFMLPVSTPPNSIAFSTGHLLVK 540  
Db 481 SNTATIIIFLPVLAELAIRLHVHPLYLMIPTGVSCSYAFMLPVSTPPNSIAFSTGHLLVK 540  
QY 541 DMVRTGLLNLNGVLLLSLAMNTWAQAIIFQLGTFPDPWANTHAANTVLTALPPALTNNVTQTL 600  
Db 541 DMVRTGLLNLNGVLLLSLAMNTWAQAIIFQLGTFPDPWANTHAANTVLTALPPALTNNVTQTL 600

RESULT 3  
US-10-167-994-4  
; Sequence 4, Application US/10167994  
; Publication No. US20030082647A1  
; GENERAL INFORMATION:  
; APPLICANT: Reenan, Robert A.  
; APPLICANT: Rogina, Blanka  
; APPLICANT: Helfand, Stephen L.  
; TITLE OF INVENTION: TRANSPORTER PROTEIN  
; FILE REFERENCE: 13407-013001  
; CURRENT APPLICATION NUMBER: US/10/167,994  
; CURRENT FILING DATE: 2002-06-12  
; PRIOR APPLICATION NUMBER: US 60/255,013  
; PRIOR FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: US 10/017,479  
; PRIOR FILING DATE: 2001-12-12  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 602  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-167-994-4

Query Match 87.0%; Score 2691; DB 14; Length 602;  
Best Local Similarity 85.9%; Pred. No. 7.6e-237;  
Matches 517; Conservative 41; Mismatches 42; Indels 2; Gaps 2;  
QY 1 MAALAAKKWWSARRLAVLLVPLALPILFALPPKGRCLYVILLMAVWCTEALPLS 60  
Db 1 MAALAAKKWWSARRLAVLLVPLALPILFALPPKGRCLYVILLMAVWCTEALPLS 60  
QY 61 VTALLPILVFPFMGILPNSKVCPOYFLDTNFIPLSLMASAIEBRLHRRILAKVLMV 120  
Db 61 VTALLPILVFPFMGILPNSKVCPOYFLDTNFIPLSLMASAIEBRLHRRILAKVLMV 120  
QY 121 GVQPARLILGMVTTSFSLMNLSTASTAMMLPIASAILKSLFGQDTRKOLPREGDST 180  
Db 121 GVQPARLILGMVTTSFSLMNLSTASTAMMLPIASAILKSLFGQDTRKOLPREGDST 180  
QY 181 AAVRNGRLTVPTQMQLASSEG-CHAEVDVAPLELPDPS-KEEHRNINWKGFLISIPY 238  
Db 181 AAVRNGRLTVPTQMQLASSEG-CHAEVDVAPLELPDPS-KEEHRNINWKGFLISIPY 238  
QY 239 SASIGGTATLTGTAPNLILGQKFFPQCDVNVFGSWFIFAPLMLLFLVGLWISFL 298  
Db 241 SASIGGTATLTGTAPNLILGQKFFPQCDVNVFGSWFIFAPLMLLFLVGLWISFL 300  
QY 299 YGMSWRGRKNSKLDQVAEDKAKAVIQEENLPIKFAEQAVFIFLCLFALLFSRD 358  
Db 301 YGGLSFRGRKNSKIRTNADRARAVIREYQNLGPIKFAEQAVFIFLCLFALLFSRD 360  
QY 359 PKFIFGWSLFPAGFVSDAVTGAIVTILFFPPSQPSLKWDFDKAPNSETTEPLLSWK 418  
Db 361 PKFIFGWSLFPAGFVSDAVTGAIVTILFFPPSQPSLKWDFDKAPNSETTEPLLSWK 420  
QY 419 AQETVPWIIILLGGGFAMAKCESGLSAGWIGQLHPLHVPBLLAVLLITVVIAFTTE 478  
Db 421 AQETVPWIIILLGGGFAMAKCESGLSAGWIGQLHPLHVPBLLAVLLITVVIAFTTE 480  
QY 479 FASNTATIIIFLPVLAELAIRLHVHPLYLMIPTGVSCSYAFMLPVSTPPNSIAFSTGHLL 538

Db 481 FASNTATIIIFLPVLAELAIRLHVHPLYLMIPTGVSCSYAFMLPVSTPPNSIAFSGHLL 540  
QY 539 VKDMVRTGLLNLNGVLLLSLAMNTWAQAIIFQLGTFPDPWANTHAANTVLTALPPALTNNVTQ 598  
Db 541 VKDMVRTGLLNLNGVLLLSLAMNTWAQAIIFQLGTFPDPWADWYSVNVVTALPPTLANDTFR 600  
QY 599 TL 600  
Db 601 TL 602

RESULT 4  
US-10-108-260A-3862  
; Sequence 3862, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3862  
; LENGTH: 520  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-3862

Query Match 75.2%; Score 2326.5; DB 15; Length 520;  
Best Local Similarity 81.4%; Pred. No. 1.3e-203;  
Matches 452; Conservative 36; Mismatches 30; Indels 37; Gaps 3;  
QY 48 MAVWCTEALPLSVTALLPIILFPPMGILPNSKVCPOYFLDTNFIPLSLMASAIEERN 107  
Db 1 MAVWCTEALPLSVTALLPIILFPPMGILPNSKVCPOYFLDTNFIPLSLMASAIEERN 60  
QY 108 LHRRIALKVLMVGVQPARLILGMVTTSFSLMNLSTASTAMMLPIASAILKSLFGQD 167  
Db 61 LHRRIALKVLMVGVQPARLILGMVTTSFSLMNLSTASTAMMLPIASAILKSLFGQD 120  
QY 168 TRKOLPREGDSTAAVRNGRLTVPTQMQLASSEG-CHAEVDVAPLELPDPS-KEEHR 225  
Db 121 VRKDPQSEBNTAAVRNGRLTVPTQMQLASSEG-CHAEVDVAPLELPDPS-KEEHR 180  
QY 226 RNIWKGFLISIPYASIGGTATLTGTAPNLILGQKFFPQCDVNVFGSWFIFAPLML 285  
Db 181 RNIWKGFLISIPYASIGGTATLTGTAPNLILGQKFFPQCDVNVFGSWFIFAPLML 240  
QY 286 LFLVGLWISFLYGGMSWRGRKNSKLDQVAEDKAKAVIQEENLPIKFAEQAVFI 345  
Db 241 LFLVGLWISFLYGGMSWRGRKNSKLDQVAEDKAKAVIQEENLPIKFAEQAVFI 292  
QY 346 LFLVGLWISFLYGGMSWRGRKNSKLDQVAEDKAKAVIQEENLPIKFAEQAVFI 405  
Db 293 LFLVGLWISFLYGGMSWRGRKNSKLDQVAEDKAKAVIQEENLPIKFAEQAVFI 325  
QY 406 PNSTEPLLSWKKAQETVPWIIILLGGGFAMAKCESGLSAGWIGQLHPLHVPBLLA 465  
Db 326 PNSTEPLLSWKKAQETVPWIIILLGGGFAMAKCESGLSAGWIGQLHPLHVPBLLA 385  
QY 466 VLLITVVIAFTTEPASNTATIIIFLPVLAELAIRLHVHPLYLMIPTGVSCSYAFMLPVST 525  
Db 386 VLLITVVIAFTTEPASNTATIIIFLPVLAELAIRLHVHPLYLMIPTGVSCSYAFMLPVST 445  
QY 526 PPSNIAFSTGHLLVKDMVRTGLLNLNGVLLLSLAMNTWAQAIIFQLGTFPDPWANTHAAN 585  
Db 446 PPSNIAFSTGHLLVKDMVRTGLLNLNGVLLLSLAMNTWAQAIIFQLGTFPDPWADWYSVNV 505  
QY 586 TALPPALTNNVTQTL 600  
Db 506 TALPPALTNNVTQTL 600



APPLICANT: Spaderma, Steven K.  
APPLICANT: Catterton, Elina  
APPLICANT: Leite, Mario W.  
APPLICANT: Zhong, Haihong  
APPLICANT: Alsobrook, John P.  
APPLICANT: Lepley, Denise M.  
APPLICANT: Rieger, Daniel K.  
APPLICANT: Burgess, Catherine E.  
TITLE OF INVENTION: No. US20040043302A1el Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-290C  
CURRENT APPLICATION NUMBER: US/10/092,900A  
CURRENT FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: USSN 60/274,322  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/283,675  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: USSN 60/338,092  
PRIOR FILING DATE: 2001-12-03  
PRIOR APPLICATION NUMBER: USSN 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/274,191  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/325,681  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: USSN 60/304,354  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: USSN 60/279,995  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: USSN 60/294,899  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: USSN 60/287,424  
PRIOR FILING DATE: 2001-04-30  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 768  
SEQ ID NO 48  
LENGTH: 616  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-092-900A-48

Query Match 45.7%; Score 1412.5; DB 15; Length 616;  
Best Local Similarity 46.6%; Pred. No. 6.4e-120;  
Matches 292; Conservative 102; Mismatches 171; Indels 61; Gaps 13;

Qy 4 LAALAKKWSARRLLVLLVPLALLPILFALPKBGRCLYVILLMAVYVTEALPLSVTA 63  
Db 1 MASALSYVSKFSPVILFVTPVLLPILVILPAKVCSCAYVILLMAVYVTEVPLAVTS 60

Qy 64 LLPILPFPMGILPSSKVCQYFDTNPLFLSLGLIMASAEERNLHRIALKVLMVGQV 123  
Db 61 LMPVLLPFLQILDSRQVCQYMKDTNNMLFLGLLIVAVAVERNLHRIALKVLMVGAK 120

Qy 124 PARLILGMVTTSLSMWLSNTASTAMMLPIASAILK-----SLFGQDRDKDL 172  
Db 121 PARMLGFMGTALLSMWISNTATTAMVPIVEALQOMEATSATAGLEGQGTNNL 180

Qy 173 PREGEDSTAAVRGNGLRTVPTMQLASSSEGHAEVPEAPLE-----LPDD 218

Db 412 TLLFVPSQPK-----FNFRSQTBEKSPVLIAPPPDLWDKVTQEKVPWGI VLLGGGFA 467  
Qy 437 MAKGCEBSGLSAGIQLHLEHPVPLLAIVLITVIAFFTEFASNTATIIIFLPVLAEL 496  
Db 468 LAKGEASGLSVWKGQKWEPLHAPPAITLILSLVAVFTECTSNVATITLFLPIFASM 527  
Qy 497 AIRLHVPLYLIMIPGTVSCSVAFMLPVSTPPNSIAFSTGHLLVQDMVTRTGLMLNMGVLL 556  
Db 528 SRSLGNPLYLIMPLCTLSASFAPMLPVATPBNVAVTYGHLKVDWMTGTVMNIIGVFC 587  
Qy 557 LSLAMTWQAQIFOLGTTPDWMAN-TH 581  
Db 588 VFLAVNTWGRAIFDLDFHFPDWMANVTH 613

RESULT 7  
US-10-403-161-68  
; Sequence 68, Application US/10403161  
; Publication No. US20040043930A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-573C  
; CURRENT APPLICATION NUMBER: US/10/403,161  
; PRIOR APPLICATION NUMBER: 2003-03-31  
; PRIOR FILING DATE: 2003-03-31  
; PRIOR APPLICATION NUMBER: 60/370349  
; PRIOR FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: 60/384543  
; PRIOR FILING DATE: 2002-05-30  
; PRIOR APPLICATION NUMBER: 60/370969  
; PRIOR FILING DATE: 2002-04-08  
; PRIOR APPLICATION NUMBER: 60/403748  
; PRIOR FILING DATE: 2002-08-15  
; PRIOR APPLICATION NUMBER: 60/372019  
; PRIOR FILING DATE: 2002-04-12  
; PRIOR APPLICATION NUMBER: 60/374379  
; PRIOR FILING DATE: 2002-04-22  
; PRIOR APPLICATION NUMBER: 09/779679  
; PRIOR FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/181045  
; PRIOR FILING DATE: 2000-02-08  
; PRIOR APPLICATION NUMBER: 10/055877  
; PRIOR FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: 60/262892  
; PRIOR FILING DATE: 2001-01-19  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 173  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 68  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-403-161-68

Query Match 45.7%; Score 1412.5; DB 15; Length 616;  
Best Local Similarity 46.6%; Pred. No. 6.4e-120;  
Matches 292; Conservative 102; Mismatches 171; Indels 61; Gaps 13;

Qy 4 LAALAKKWSARRLLVLLVPLALLPILFALPKBGRCLYVILLMAVYVTEALPLSVTA 63  
Db 1 MASALSYVSKFSPVILFVTPVLLPILVILPAKVCSCAYVILLMAVYVTEVPLAVTS 60

Qy 64 LLPILPFPMGILPSSKVCQYFDTNPLFLSLGLIMASAEERNLHRIALKVLMVGQV 123  
Db 61 LMPVLLPFLQILDSRQVCQYMKDTNNMLFLGLLIVAVAVERNLHRIALKVLMVGAK 120

Qy 124 PARLILGMVTTSLSMWLSNTASTAMMLPIASAILK-----SLFGQDRDKDL 172  
Db 121 PARMLGFMGTALLSMWISNTATTAMVPIVEALQOMEATSATAGLEGQGTNNL 180

Qy 173 PREGEDSTAAVRGNGLRTVPTMQLASSSEGHAEVPEAPLE-----LPDD 218

Db 181 NALBDDTVKAVLGG--KCVAIITVYKVKVKQLINMLTLPKLEKQEQDQLGPIRPQD 238  
QY 219 S---KESEHRRNIWKGFLISIPYSASIGGTATLTGTAPNLLILGOLKSPFQOC-DVNVFG 274  
Db 239 SAQCOEDQERKRLCKAWTLCICYAASIGGTATLTGTGNVLLGOMNELFPDSKDLVNFPA 298  
QY 275 SWFTFAPPLMLLFLVGLWMLISFLYGGMSW-RGW-----RKNKSKLDQVAEDKAKAVIOE 328  
Db 299 SWFAFAPNMLVMLLPAWMLQFYMFSSPKSGCGLESKQNEK-----AALKVLQOE 351  
QY 329 EFQNLGPIKFAQAVFLFCFLFAILLPSRDPKPIGWASL-FAPG-----FVSDAVTGVAVI 384  
Db 352 EYRKLGLSFAEINVLICFFLLVILWFSDPGFPGWLTVAWVEGETKSVSDATVAIFVA 411  
QY 385 TILFPFSPQPSLKWDFEAPNSETSE-----PLLSMKKAQETVPWNIILLLGGGFA 436  
Db 412 TLLFIVSPQPK----FNFSQTEEGSPVLIIAPPPLLDWKVTQKVPWGLVILLGGGFA 467  
QY 437 MAKCEBSGLSAMTGGQLHPLHVPPLAVLLIIVIAFFTEFASNTATIIIFLPVLAEL 496  
Db 468 LAKGSEASGLSVWVGKQWELHVPAAITLILSLVAVFTECTSNVATTTFLPIFASM 527  
QY 497 AIRLHVPLIYMIPTGTVSCSVAFMLPVSTPNSIATFSTGHLVVKDMVRTGLLMLNMGVLL 556  
Db 528 SRSIGLNPXYIMLPCTLSASFAFMLPVATPPNAIVFTYGHKLVADMVKTGVMNIIGVFC 587  
QY 557 LSLAMNTWAQAI FOLGTFPDWAN-TH 581  
Db 588 VFLAVNTWGRAIFDLDPHDHPPDANVTH 613

RESULT 8  
US-10-173-519-2  
; Sequence 2, Application US/10173519  
; Publication No. US20020193582A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; TITLE OF INVENTION: 69624, A Human Transporter Family Member  
; TITLE OF INVENTION: and Uses Therefor  
; FILE REFERENCE: MF01-098PIRNM  
; CURRENT APPLICATION NUMBER: US/10173,519  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: 60/298,970  
; PRIOR FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 568  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-173-519-2  
Query Match 45.6%; Score 1411.5; DB 13; Length 568;  
Best Local Similarity 47.1%; Pred. No. 7.1e-120;  
Matches 286; Conservative 100; Mismatches 150; Indels 71; Gaps 13;  
QY 4 LAALAKVWSARRLLVLLVPLLPILFALPKKEGRCCLYVILLMAVYVCTEALPLSVTA 63  
Db 1 MASALSVSXFKSPSVILFVTPLLPLVILPMPAKFVRCACAVIILMAYVCTEVIPLAVTS 60  
QY 64 LLPILPFPFGIILPSSKVCQYFIDTNTFLSLGLIMASAEERNLHRRALKVLMVGVQ 123  
Db 61 LMPVLLPFLQILDQRVCYQVKDTNMLFLGLLIVAVAVERNLHRRALKVLMVGVQ 120  
QY 124 PARILGMVMTTSLSMWLNSTASTAMPLPIASAILKSLFGQDRDKDLPREGEDSTAAV 183  
Db 121 PARMLGMVMTTSLSMWLNSTASTAMPLPIASAILKSLFGQDRDKDLPREGEDSTAAV 183  
QY 184 RGNGLRVTPTMQPLASSEGHAEDVEAPLELPPD-----SKEEHRNRNIWKGFL 233  
Db 166 TEAGLELV-----DKGKAK-----ELPGSQVIFGPTLGGQEDQERKRLCKAWT 209

QY 234 ISIPYSASIGGTATLTGTAPNLLILGOLKSPFQOC-DVNVFGSWFIFAPPLMLLFLVGV 292  
Db 210 LCICVYASIGGTATLTGTGNVLLGOMNELFPDSKDLVNFASWFAFAPNMLVMLLPAW 269  
QY 293 LMSIFLYGGMSW-RGW-----RKNKSKLDQVAEDKAKAVIOEFGNLPKFAEQAVFIL 346  
Db 270 LMLQFYMYRNFNPKSGCGLESKQNEK-----AALKVLQOEYRKLGLSFAEINVLIC 322  
QY 347 FCLFAILLPSRDPKPIGWASL-FAPG-----FVSDAVTGVAVIIVTILFPFSPQPSLKWDF 402  
Db 323 FELLVILWFSDPGFPGWLTVAWVEGETKYVSDATVAIFVATLFLFIVSPQPK----FN 378  
QY 403 FKAPNSETSE-----PLLSMKKAQETVPWNIILLLGGGFAFAMAKCEBSGLSAMIGQQLH 455  
Db 379 FRSQTEERKTFYPPPLLDWKVTQKVPWGLVILLGGGFAKSEASGLSVWVGKQME 438  
QY 456 PLEHVPPLAVLLIIVIAFFTEFASNTATIIIFLPVLAELAIRLHVPLIYMIPTGTVSC 515  
Db 439 PLHAVPPAAITLILSLVAVFTECTSNVATTTFLPIFASMSRSIGLNPYIMLPCTLSA 498  
QY 516 SYAFMLPVSTPNSIATFSTGHLVVKDMVRTGLLMLNMGVLLSLAMNTWAQAI FOLGTF 575  
Db 499 SFAFMLPVATPPNAIVFTYGHKLVADMVKTGVMNIIGVFCVFLAVNTWGRAIFDLDPHD 558  
QY 576 DWAN-TH 581  
Db 559 DWANVTH 565  
RESULT 9  
US-10-403-161-72  
; Sequence 72, Application US/10403161  
; Publication No. US20040043930A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-573C  
; CURRENT APPLICATION NUMBER: US/10/403,161  
; CURRENT FILING DATE: 2003-03-31  
; PRIOR APPLICATION NUMBER: 60/370349  
; PRIOR FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: 60/384543  
; PRIOR FILING DATE: 2002-05-30  
; PRIOR APPLICATION NUMBER: 60/370969  
; PRIOR FILING DATE: 2002-04-08  
; PRIOR APPLICATION NUMBER: 60/403748  
; PRIOR FILING DATE: 2002-08-15  
; PRIOR APPLICATION NUMBER: 60/372019  
; PRIOR FILING DATE: 2002-04-12  
; PRIOR APPLICATION NUMBER: 60/374379  
; PRIOR FILING DATE: 2002-04-22  
; PRIOR APPLICATION NUMBER: 09/779679  
; PRIOR FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/181045  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 10/055877  
; PRIOR FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: 60/262892  
; PRIOR FILING DATE: 2001-01-19  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 173  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 72  
; LENGTH: 568  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-403-161-72  
Query Match 45.6%; Score 1411.5; DB 15; Length 568;  
Best Local Similarity 47.1%; Pred. No. 7.1e-120;  
Matches 286; Conservative 100; Mismatches 150; Indels 71; Gaps 13;  
QY 4 LAALAKVWSARRLLVLLVPLLPILFALPKKEGRCCLYVILLMAVYVCTEALPLSVTA 63

Db	1	MASALSYVSKPKSFVILFVTPULLPLVILMPAKFVRCAYVILMAIYWCTEVIPLAVTS	60
Qy	64	LLPIILPPFMGILPSSKVCPOYFLDTNLFSLGLIMASAEERNLHRRIALKVLMLVGQY	123
Db	61	LMPVLLPFLFOILDSRQVCVQYMKDTNMLFLGGLIVAVAVERNLHKRIALRTLWVGAK	120
Qy	124	PARLILGNMVTTSPLSMWLSNTASTAMMLPIASAILKSLFGQDRTKDLPRGEDOSTAAV	183
Db	121	PARLMLGPMGVALLUSMISINTATTAMMVPIVEAILQOM-----BATSAA	165
Qy	184	RGNGLRTVPTMQFVLASSEGHAEDVEAPLELPDD-----SKBEHRRNIWKGL	233
Db	166	TEAGLELV-----DKGKAK-----ELPGSQVIFEGPTLGGQEDQERKLCKAMT	209
Qy	234	ISIPYSASIGGTATLTGTAPNLILILGOLKSPFPQC-DVNFQSGWIFAPPLMLLFLVGN	292
Db	210	LCICYAASIGGTATLTGTGPNVVLGQMNELFPDSKDLVNFASFAPAPPNLMLLFAW	269
Qy	293	LWISFLYGMGSW-RGW-----RKKNKSLQDVAEDKAKAVIOBFONLGPKEFAEQAVFL	346
Db	270	LWLQFVYMRPFKSKSGCGLESKKVEK-----NALKVLQBEYRKGLPGLSPAEINVLIC	322
Qy	347	FCLFAILLFSRDPKFIPOWASL-FAPG---FVSDAVTGVAVITIIILFPFPQSPKSLKWNPD	402
Db	323	FFLVLILWFSDPGFMGPGLTWAVWGEGETKYVSDATVAIFATLLFIVPSQKPK---FN	378
Qy	403	FKAPNSTE-----PILLSKKAQETVPWNIIILLGGGFAMAKGCBSSGLSAGIGQLH	455
Db	379	FRSQTEERKTPFPVPPPLLDKMKVTQEKVPWGIIVLLGGGFALAKGSEAGLSVMNKGQNE	438
Qy	456	PLEHVPPLLAVLLITVWIAFTTEFASNTATIIIFLPVLAEIALRUHVHPLYLMIPGTVSC	515
Db	439	PLHAVPPAAIPLIILSLVAVFECTSNVATTTFLFPIFASMSRSIGLNPYIMLPCTLSA	498
Qy	516	SYAFMLPVSTPPNSIARSTGHLLVKDMVTRTGLLNNLMGVLLLSLAWNTWQAQIPLQGTGP	575
Db	499	SFAFMLPVATPDPAIVFTYGHKLKQADMVKTIQVIMNIIIGVCFVLAVNTWGRAIFOLDHFP	558
Qy	576	DWAN-TH 581	
Db	559	DWANVTH 565	

```

RESULT 11
US-10-718-359-6
; Sequence 6, Application US/10718359
; Publication No. US20050095240A1
; GENERAL INFORMATION:
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.
; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION
; FILE REFERENCE: 275.0008 0101
; CURRENT APPLICATION NUMBER: US/10718,359
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/428,469
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 60/459,441
; PRIOR FILING DATE: 2003-04-01
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 568
; TYPE: PRT
; ORGANISM: human NACT
US-10-718-359-6

```

```

RESULT 10
US-10-490-080-1
; Sequence 1, Application US/10490080
; Publication No. US20040253597A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Novel Protein and its DNA
; FILE REFERENCE: P02-0109PCT
; CURRENT APPLICATION NUMBER: US/10/490.080
; CURRENT FILING DATE: 2004-03-17
; PRIORITY APPLICATION NUMBER: JP 2001-281992
; PRIORITY FILING DATE: 2001-09-17
; PRIORITY APPLICATION NUMBER: JP 2001-306873
; PRIORITY FILING DATE: 2001-10-02
; PRIORITY APPLICATION NUMBER: JP 2002-113279
; PRIORITY FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 42
; SEQ ID NO 1
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Human
US-10-490-080-1

```

**Query Match**      45.6%; Score 1411.5; DB 16; Length 568;  
**Best Local Similarity**    47.1%; Pred.No. 7.1e-120;  
**Matches**         286; Conservative 100; Mismatches 150; Indels    71; Gaps    13  
  
**Qy**                  4 LALAKKVGARRLLVLLLPLALLPILFAPPKEGRCLYVLIMAVVTCTEALPLSVTA 63  
                  : :   :   :   :   :   :   :   :   :   :   :   :   :  
                  : :   :   :   :   :   :   :   :   :   :   :   :   :



TYPE: PRT  
ORGANISM: Xenopus laevis  
US-09-729-094-4

Query Match 45.4%; Score 1405.5; DB 9; Length 619;  
Best Local Similarity 45.1%; Pred. No. 2.8e-119;  
Matches 280; Conservative 122; Mismatches 170; Indels 49; Gaps 11;

QY 4 LAALAKKWSARRLLVLLVPLALPILPALPPKGRCLYVILLMAVYVCTEALPLSVTA 63  
Db 1 MVSIGKWLARNYFIIFLPLPLPLVPLVPTKEASCGFVIVMALFWCTEALPLAVTA 60

QY 64 LPILPFPFMGILPSSKVCPOVFLDTNLFSLGLMASAIEERNLHRRALVKVLMVGVO 123  
Db 61 LFPVLLFPFMGIMDSTAVCSQYLKDTNMLFICGLLVAISVEKNLHRRALVKVLMVGVO 120

QY 124 PARLILGMVMTTSLFSLMWSNTASTAMMLPIASAILKSLF-----GORDTRKDL 172  
Db 121 PALLLLGFVAVTAFSLMWSNTATTAMMIPIAQAVMEQLHSSEKGVDERVEGNSNTQKV 180

QY 173 PREGEDSTAARVGNLRTVPTMQFLASSEGHAEDVE-----APLEL-PD 217  
Db 181 NGMENDMTESVMPSGMALAIDNTVATENEGFEIOEKSTKDPSPKQSGIGPVIIEPE 240

QY 218 DSKEEHR-----NIWKGFLISIPYSASIGTATLTGTAPNLILLGQKSPFPQ-CDVV 271  
Db 241 DEKQTEEKQKHEKHLKICKGMSLCVCSASIGGIATLTGTTPLNVMKGOMDELFPENNII 300

QY 272 NFGSWFIAPPLMLFLVLGWLWISFLYGMWSRGRKNSKLQDVAEDKAK-----AVI 326  
Db 301 NFASWFGFAPPTMLVLLSALSWLQFIYLVNF-----KNFGCGGNASQKQEKRAFRVI 356

QY 327 QEEFNQGLPIFAEQVFIPLFCFAILLFSRDPKFIQWASL-FAPG---FVSDAVTGA 382  
Db 357 SGEHKLGSMTPFAISVLVPLILLVLLFTFREPFGMPGWATISFNKGGKEMVTDATVAF 416

QY 383 IVTILFPFPPSKPKLWDFDK---APNSETP-LLSWKKAQETVPMNIILLGGGFAMA 438  
Db 417 VSLMFFPFPSPKPKLWDFDK---APNSETP-LLSWKKAQETVPMNIILLGGGFAMA 476

QY 439 KGESGLSAMIGGQLHLEHVPPLVLLVITVIAFTEFASNTATIIIFLPVLAELAI 498  
Db 477 KGESGLSLWLGKLTPLQSIIPPAIALICLLVATFECTSNVATTTFLFILASMAK 536

QY 499 RLHVHPLYMIPGTVSCSYAFMLPVSTPPNSIAFSTGHLLVKDMVRTGLLMLMGVLLS 558  
Db 537 AIQLNPLYMLPCTLSASLAFMLPVATPPNAIFSYGQLKVIDMAKAGLLNIGLVLTIT 596

QY 559 LAMNTWAQAIIFOLGTFPPDAN 579  
Db 597 LAINSWGFMENLGTFFPSWAN 617

RESULT 14  
US-10-435-631-4  
; Sequence 44, Application US/10435631  
; Publication No. US20030186381A1  
; GENERAL INFORMATION:  
; APPLICANT: CHATURVEDI, Kabir et al  
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CLO00662CON  
; CURRENT APPLICATION NUMBER: US/10/435,631  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 619  
; TYPE: PRT  
; ORGANISM: Xenopus laevis  
US-10-435-631-4

Query Match 45.4%; Score 1405.5; DB 14; Length 619;  
Best Local Similarity 45.1%; Pred. No. 2.8e-119;  
Matches 280; Conservative 122; Mismatches 170; Indels 49; Gaps 11;

QY 4 LAALAKKWSARRLLVLLVPLALPILPALPPKGRCLYVILLMAVYVCTEALPLSVTA 63  
Db 1 MVSIGKWLARNYFIIFLPLPLPLVPLVPTKEASCGFVIVMALFWCTEALPLAVTA 60

QY 64 LPILPFPFMGILPSSKVCPOVFLDTNLFSLGLMASAIEERNLHRRALVKVLMVGVO 123  
Db 61 LFPVLLFPFMGIMDSTAVCSQYLKDTNMLFICGLLVAISVEKNLHRRALVKVLMVGVO 120

QY 124 PARLILGMVMTTSLFSLMWSNTASTAMMLPIASAILKSLF-----GORDTRKDL 172  
Db 121 PALLLLGFVAVTAFSLMWSNTATTAMMIPIAQAVMEQLHSSEKGVDERVEGNSNTQKV 180

QY 173 PREGEDSTAARVGNLRTVPTMQFLASSEGHAEDVE-----APLEL-PD 217  
Db 181 NGMENDMTESVMPSGMALAIDNTVATENEGFEIOEKSTKDPSPKQSGIGPVIIEPE 240

QY 218 DSKEEHR-----NIWKGFLISIPYSASIGTATLTGTAPNLILLGQKSPFPQ-CDVV 271  
Db 241 DEKQTEEKQKHEKHLKICKGMSLCVCSASIGGIATLTGTTPLNVMKGOMDELFPENNII 300

QY 272 NFGSWFIAPPLMLFLVLGWLWISFLYGMWSRGRKNSKLQDVAEDKAK-----AVI 326  
Db 301 NFASWFGFAPPTMLVLLSALSWLQFIYLVNF-----KNFGCGGNASQKQEKRAFRVI 356

QY 327 QEEFNQGLPIFAEQVFIPLFCFAILLFSRDPKFIQWASL-FAPG---FVSDAVTGA 382  
Db 357 SGEHKLGSMTPFAISVLVPLILLVLLFTFREPFGMPGWATISFNKGGKEMVTDATVAF 416

QY 383 IVTILFPFPPSKPKLWDFDK---APNSETP-LLSWKKAQETVPMNIILLGGGFAMA 438  
Db 417 VSLMFFPFPSPKPKLWDFDK---APNSETP-LLSWKKAQETVPMNIILLGGGFAMA 476

QY 439 KGESGLSAMIGGQLHLEHVPPLVLLVITVIAFTEFASNTATIIIFLPVLAELAI 498  
Db 477 KGESGLSLWLGKLTPLQSIIPPAIALICLLVATFECTSNVATTTFLFILASMAK 536

QY 499 RLHVHPLYMIPGTVSCSYAFMLPVSTPPNSIAFSTGHLLVKDMVRTGLLMLMGVLLS 558  
Db 537 AIQLNPLYMLPCTLSASLAFMLPVATPPNAIFSYGQLKVIDMAKAGLLNIGLVLTIT 596

QY 559 LAMNTWAQAIIFOLGTFPPDAN 579  
Db 597 LAINSWGFMENLGTFFPSWAN 617

RESULT 15  
US-10-092-900A-44  
; Sequence 44, Application US/10092900A  
; Publication No. US20040043382A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Spvtek, Kimberly A.  
; APPLICANT: Shenov, Sureeh G.  
; APPLICANT: Taupier Jr., Raymond J.  
; APPLICANT: Pena, Carol E.A.  
; APPLICANT: Li, Li  
; APPLICANT: Zehusen, Bryan D.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Gorman, Linda  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Gangolli, Esha A.  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Guo, Xiaojia Sasha T.  
; APPLICANT: Tchernev., Velizar T.  
; APPLICANT: Fernandes, Elma R.  
; APPLICANT: Casman, Stacie J.

APPLICANT: Malyankar, Uriel M.  
APPLICANT: Gerlach, Valerie  
APPLICANT: Liu, Yi  
APPLICANT: Anderson, David W.  
APPLICANT: Spaderna, Steven K.  
APPLICANT: Catterton, Elina  
APPLICANT: Leite, Mario W.  
APPLICANT: Zhong, Haihong  
APPLICANT: Alsobrook, John P.  
APPLICANT: Lepley, Denise M.  
APPLICANT: Rieger, Daniel K.  
APPLICANT: Burgess, Catherine E.  
TITLE OF INVENTION: No. US20040043382A1e1 Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-290C  
CURRENT APPLICATION NUMBER: US/10/092,900A  
CURRENT FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: USSN 60/274,322  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/283,675  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: USSN 60/338,092  
PRIOR FILING DATE: 2001-12-03  
PRIOR APPLICATION NUMBER: USSN 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/274,191  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: USSN 60/325,681  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: USSN 60/304,354  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: USSN 60/279,995  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: USSN 60/294,899  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: USSN 60/287,424  
PRIOR FILING DATE: 2001-04-30  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 768  
SEQ ID NO 44  
LENGTH: 568  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-092-900A-44

Query Match 45.3%; Score 1400.5; DB 15; Length 568;  
Best Local Similarity 47.0%; Pred. No. 7.2e-119;  
Matches 286; Conservative 100; Mismatches 149; Indels 73; Gaps 14;  
QY 4 LAALAKKVSARRLLVLLVPLALLPILFALPPKRGCLYVILLMAVYWCTEALPLSVTA 63  
DB 1 MASALSYSKFSFVILFVTPLLLPLVILMPAKVS-CAYVILMAIYWCTEVIPLAVTS 59  
QY 64 LLPILPPFMGILPSSKVCQYFLDTNFTLSGLIMASAIERNLHRIALKVLMVGWQ 123  
DB 60 LMPVLLPFLQILDQRVCQYKMDTNMLFLGGLIVAVAVERNWLHRIALRTLLMWGAK 119  
QY 124 PARILGCMVYTSFLSNLSTANMLPLASAILKSLFGQDRDKLPRGEDSTAAV 183  
DB 120 PARMLGMFGVTALLSWISNTATTAMVPIVEAILQOM-----EATSA 164  
QY 184 RGNGLRVTPTMQFLASSEGHAEDVEAPLELPD-----SKEEHRNRNIWKGL 233  
DB 165 TEAGLELV-----DKGKAK-----ELPGSQVIFSGPTLGQDQERKRLCKAMT 208  
QY 234 ISIPYSASIGGTATLTGTAPNLILLQGLKSFPPQC-DVNFPGSWFIFAFPLMLLFLVGW 292  
DB 209 LCICYAASIGGTATLTGTGNVLLGQNELFPDSKDLNVNFASFAPAFENMLVMLFAW 268  
QY 293 LWISFLYGGMSW-RGW-----RKNKSKLDQVAEDKAKAVIQEBFONLGPICFAEQAVFIL 346  
DB 269 LWQFVYVNFSSFKKSWCGGLESKNEK-----AALKVIOEYRKLPGLSFAEINVLC 321  
QY 347 FCLFAILLFSRDPKFIPEGWASL-FAPG---FVSDAVTGAIVTILFPFPSPKPSLKWTFD 402

Db 322 FELLVILWFSRDPGFMGLTVAWVEGETKQYVSDATVAIFVATLLFIVPSQPK-----FN 377  
QY 403 EKAPNSET-----PLLSWKKAQETVPKNIILLGGGFAMAKGCEESGLSAMIGQL 454  
Db 378 FRSQTEGKSPVLIAPPPLLDWKVTQEKVPMGIVLLGGGFALAKGSEASGLSVWVKQM 437  
QY 455 HPLEHVPPLLAVALLITVVIATFEPASNTATIIIFLPVLAELAIRLHVHPLYLMIPTVS 514  
Db 438 EPLHAVPPAAITLIISLVAVFECTSNVATTTLELPFASMSRISGLNPLYIMLPCTLS 497  
QY 515 CSYAFMLPVSTPPNSIAFSTGHLLVKDVMRTGLLMLNMGVLLLSLAMNTWAOAIFQLGTF 574  
Db 498 ASFAFMPLPVATPPNAIVFTYGLKQVADMVKTGVIIMNIIGVFCVFLAVNTWGRAIFDLDF 557  
QY 575 PDWAN-TH 581  
Db 556 PDWANVTH 565

Search completed: June 30, 2005, 09:32:52  
Job time : 79.8579 secs